# Insights into Protein-Ligand Molecular Recognition: Thermodynamic, Kinetic and Structural Characterization of Inhibitor Binding to Aldose Reductase and Carbonic Anhydrase II

Dissertation

zur Erlangung des Doktorgrades
der Naturwissenschaften
(Dr. rer. nat.)

dem
Fachbereich Pharmazie der
Philipps - Universität Marburg
vorgelegt

von

Chris Rechlin

aus

Flensburg

Marburg/Lahn 2015

Erstgutachter: **Prof. Dr. Gerhard Klebe**

Zweitgutachter: **Prof. Dr. Andreas Heine**

Eingereicht am **17.08.2015**

Tag der mündlichen Prüfung am **14.10.2015**

Hochschulkennziffer: 1180

Die Untersuchungen zur vorliegenden Arbeit wurden auf Anregung von Herrn Prof. Dr. Gerhard Klebe am Institut für Pharmazeutische Chemie der Philipps-Universität Marburg in der Zeit vom Juli 2011 bis August 2015 durchgeführt.

# Abbrevations

| | |
|---|---|
| 2-h PG | 2-h plasma glucose |
| Å | Ångström (1Å= $10^{-10}$ m) |
| A1C | Glycated hemoglobin |
| $A_{600}$ | Absorption at 600 nm |
| ADME | Absorption, distribution, metabolism and excretion |
| AGE | Advanced glycation end product |
| AKR1A1 | Aldo-keto reductase family 1, member A1 (aldehyde reductase) |
| AKR1B1 | Aldo-keto reductase family 1, member B1 (aldose reductase) |
| AKR1B10 | Aldo-keto reductase family 1, member B10 |
| Ala | Alanine |
| ALR1 | Human aldehyde reductase |
| ALR2 | Human aldose reductase |
| APCI | Atmospheric pressure chemical ionisation |
| API | Active pharmaceutical ingredient |
| APS | Ammoniumpersulfate |
| Arg | Arginine |
| ARI | Aldose reductase inhibitor |
| Asn | Asparagine |
| Asp | Aspartic acid |
| BEI | Binding efficiency index |
| *B*-factor | Debye-Waller factor |
| CA | Carbonic anhydrase |
| CAII | Human carbonic anhydrase II |
| CARP | Carbonic anhydrase related proteins |
| CV | Collective Variable |
| Cys | Cysteine |
| D | Distribution ratio |
| Da | Dalton |
| DMSO | Dimethyl sulfoxide |
| DPP-4 | Dipeptidyl peptidase-4 |
| DTT | Dithiothreitol |
| *E. coli* | Escherichia coli |
| EDC | 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide |
| EDTA | Ethylene diamine tetra acetate |
| $F_c$ | Calculated structure amplitudes |
| $F_o$ | Observed structure amplitudes |
| FPG | Fasting plasma glucose |

# Abbrevations

| | |
|---|---|
| $G^0$ | Gibbs free energy |
| Gln | Glutamine |
| GLP-1 | Glucagon-like peptide-1 |
| Glu | Glutamic acid |
| Gly | Glycine |
| GSH | Reduced glutathione |
| GSSG | Oxidized gluathione |
| GST | Glutathione-S-transferase |
| h | Hour |
| $H^0$ | Enthalpy |
| H-bond | Hydogen bond |
| HEPES | 4-(2-Hydroxyethyl)-1-piperazine ethan sulfonic acid |
| His | Histidine |
| HP | Hydrophobic patch |
| $IC_{50}$ | Half maximal inhibitory concentration |
| IDD | Institute for diabetes discovery |
| Ile | Isoleucine |
| IPTG | Isopropyl-β-D-thiogalactopyranosid |
| ITC | Isothermal microcalorimetry |
| K | Kelvin |
| k | Kilo |
| $K_A$ | Association constant |
| $K_D$ | Dissociation constant |
| $k_{off}$ | Rate of binary complex dissociation |
| $k_{on}$ | Rate of ligand association |
| LE | Ligand efficiency |
| LELP | Ligand efficiency-dependent lipophilicity |
| Leu | Leucin |
| LLE | Lipophilic ligand efficiency |
| Lys | Lysine |
| M | Molarity (mol/L) |
| MD | Molecular dynamics simulation |
| MES | 2-(*N*-morpholino)ethanesulfonic acid |
| Met | Methionine |
| min | Minute |
| $NAD^+$ | Nicotinamide adenine dinucleotide |
| $NADP^+$ | Nicotinamide adenine dinucleotide phosphate |
| NADPH | Nicotinamide adenine dinucleotide phosphate, reduced |
| NHS | N-hydroxysuccinimide |

# Abbrevations

| | |
|---|---|
| NMR | Nuclear magnetic resonance |
| PAGE | Polyacrylamide gel electrophoresis |
| PBS | Phosphate buffered saline |
| PDB | Protein Data Bank |
| PEG 6000 | Polyethylene glycole 6000 |
| pH | Potentialis hydrogenii |
| Phe | Phenylalanine |
| $pK_a$ | Logarithmic acid dissociation constant |
| PKC | Protein kinase C |
| Pro | Proline |
| R | Ideal gas constant |
| R-factor | Reliability factor |
| RMSD | Root mean square determination |
| rpm | Revolution per minute |
| $S^0$ | Entropy |
| SD | Standard deviation |
| SDS | Sodium dodecyl sulfate |
| Ser | Serine |
| SFF | Stopped-flow fluorescence |
| SGLT | Sodium-glucose transport proteins |
| siRNA | Small interfering RNA |
| SPR | Surface plasmon resonance |
| T | Absolute temperature |
| $t_{HP}$ | Lifetime of the intermediate state at the hydrophobic patch domain |
| Thr | Threonine |
| TRIS | Tris(hydroxamethyl)aminomethane |
| Trp | Tryptophane |
| Tyr | Tyrosine |
| Val | Valine |
| w/v | Weight per volume |
| X-ray | Röntgen radiation |

# Table of Contents

Table of Contents

Table of Contents

# Publications Arising from this Work

Parts of this thesis were published or were prepared for submission to a scientific journal. Each chapter is preceded by an introductory remark which clarifies the contribution of the author of this thesis.

**Articles:**

Stefek, M., Soltesova Prnova, M., Majekova, M., Rechlin, C., Heine, A. and Klebe, G. (2015) Identification of Novel Aldose Reductase Inhibitors Based on Carboxymethylated Mercaptotriazinoindole Scaffold, *J. Med. Chem. 58*, 2649−2657.

Ruiz, F. X., Cousido-Siah, A., Porté, S., Domínguez, M., Crespo, I., Rechlin, C., Mitschler, A., de Lera, Á. R., Martín, M. J., de la Fuente, J. Á., Klebe, G., Parés, X., Farrés, J. and Podjarny, A. Structural Determinants of the Selectivity of 3-benzyluracil-1-acetic Acids towards Human Enzymes Aldose Reductase and AKR1B10. In preparation.

Rechlin, C., Ortmann, R., Heine, A., Eisenmann, M., Schlitzer, M. and Klebe, G. Crystallographic and Thermodynamic Characterization of Carboxylic Acids Containing Two or Three Aromatic Systems as Inhibitors of Human Aldose Reductase. In preparation.

Rechlin, C., Scheer, F., Terwesten, F., Heine, A., Diederich, W. and Klebe, G. Thermodynamic and Structural Investigation of the Opening of the Specificity Pocket of Human Aldose Reductase. In preparation.

Gaspari, R., Rechlin, C., Heine, A., Bottegoni, G., Rocchia, W., Schwarz, D., Bomke, J., Gerber, H.-D., Klebe, G. and Cavalli, A. Kinetic and Structural Insights into the Mechanism of Binding of Sulfonamides to Human Carbonic Anhydrase. In preparation.

Publications Arising from this Work

**Posters:**

Rechlin, C., Heine, H., Ortmann, R., Schlitzer, M. and Klebe, G. Thermodynamic and Structural Characterization of Bi- and Triaryl Carboxylic Acid Derivatives as Inhibitors of the Human Aldose Reductase. Biophysics in Drug Discovery, Strasbourg, France, 2013.

Rechlin, C., Terwesten, F., Scheer, F., Toth, P., Heine, A., Diederich, W. and Klebe, G. How Much Does the Opening of the Specificity pocket of the Human Aldose Reductase cost? IDD-type ligands under investigation. Developments in Protein Interaction Analysis, San Diego, CA, USA, 2014.

Rechlin, C., Terwesten, F., Scheer, F., Toth, P., Heine, A., Diederich, W. and Klebe, G. Aldose Reductase as a Target to Prevent Diabetic Complications: Investigations of the Structural and Thermodynamic Background of the Opening of the Specificity Pocket. Trends and Perspectives in Pharmaceutical Sciences- DPhG Annual Meeting, Frankfurt, Germany, 2014.

Rechlin, C., Scheer, F., Terwesten, F., Toth, P., Heine, A., Diederich, W. and Klebe, G. Protein Flexibility- Option or Frontier in Medicinal Chemistry: What Can Be Learned from the Specificity Pocket in the Human Aldose Reductase? Frontiers in Medicinal Chemistry- GDCh Annual Meeting, Marburg, Germany, 2015.

# 1 Introduction and Aims of this Work

## 1.1 Drug Design

At the beginning of the last century serendipity was often the base for the discovery of new drugs. The most famous example for such a finding is presumably the detection of the antibiotic penicillin by Alexander Flemming in 1928. Thanks to his attentive mind he noticed that *Penicillum notatum* leads to bacteria-free areas in a plate of a *Staphylococcus* culture. Such empiric observations helped to identify active ingredients. The mode of action and the target in the organism was often elucidated long after the discovery of a drug. Today the rational drug design takes the opposite approach. It starts with the identification of possibly disease-related targets in the human or animal body, a molecular structure which can interact specifically with drugs. The pharmacological modulation of the target is validated for its intended therapeutic use. X-ray crystallography is used to solve the three-dimensional structure of the target. Possible drug molecules are designed with the help of computational approaches. The most promising compounds are synthesized and tested in biological and biophysical assays. This step is further supported by protein crystallography to obtain complex structures of the ligand and its target with the aim to generate ideas for the next design cycle. Furthermore, at early stages the ADME (absorption, distribution, metabolism and excretion) parameters and possible side effects like affinity for the hERG channel are taken into account. After the *in vitro* testing identifies a promising drug candidate this molecule is tested in animal models. In the end, phase I-III trials in humans have to prove the safety and efficacy which paves the way to the approval as a drug. Altogether, the development of one new API (active pharmaceutical ingredient) costs up to 1.6 billion US Dollars. Thus, there is a strong need for the most promising lead substances to be chosen very carefully in the early stages of the development process. A profound understanding of the process of molecular recognition between target protein and ligand is therefore essential.[1-3]

## 1.2 The Significance of Thermodynamic and Kinetic Parameters for Drug Design

The affinity of a ligand to a protein is often expressed as the half-maximal inhibitory concentration ($IC_{50}$) or the dissociation constant ($K_D$). However, the sole consideration of these values has a limited validity since screening hits often possess affinities in the same order of magnitude, which makes a more detailed analysis of the biophysical parameters desirable. Therefore, the thermodynamic profile and the kinetic constants of a ligand binding to its target protein are included in the decision making process.[4]

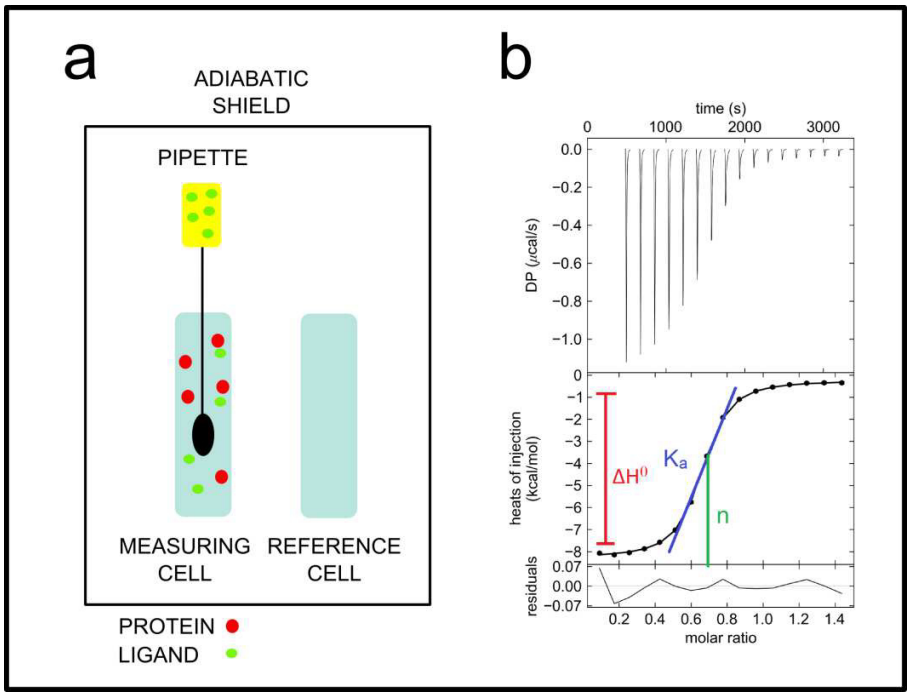



**Figure 1-1 Theoretical background of an ITC experiment** a) Schematic view of an ITC instrument. The protein solution is filled in the measuring cell and the ligand in the pipette. The reference cell is filled with water. b) Example thermogramm and fitted regression curve. The enthalpy $\Delta H^0$ corresponds to the difference of the two plateaus, the association constant $K_a$ to the slope of the curve and the stoichiometry to the x-axis value of the inflection point of the curve.

Isothermal microcalorimetry allows the dissection of the Gibbs Free Energy ($\Delta G^0$) into the enthalpy ($\Delta H^0$) and the entropy ($-T\Delta S^0$). The instrument consists of a measuring cell and a reference cell which are coated by an adiabatic shield. In the experimental set-up which was used in this work the protein was introduced in the measuring cell and the ligand in the pipette

(Figure 1-1a). Several injections of a certain volume of the ligand are performed with a time lag in between. A feedback power holds both cells on the same temperature. An exothermic reaction decreases temporarily the feedback power, as shown in Figure 1-1b. During data evaluation each of the peaks of the thermogram is integrated. The resulting value is divided by the injected amount of ligand (in mole) and plotted against the molar ratio of ligand to protein. The enthalpy $\Delta H^0$ is represented by the difference of the upper and lower plateau of the resulting fitted regression curve, the association constant $K_a$ by its slope and the stoichiometry by the x-axis value of its inflection point (Figure 1-1b). $\Delta G^0$, $\Delta H^0$ and $-T\Delta S^0$ can be calculated using equation 1-1:

$$\Delta G^0 = -RT \ln K_a = \Delta H^0 - T\Delta S^0 \qquad (1\text{-}1)$$

with R= 8.314 J $K^{-1}$ $mol^{-1}$ and absolute temperature T.[5]

The characterization of the full thermodynamic profile can provide valuable information. A favorable $\Delta H^0$ indicates a net increase in the number and strength of non-covalent bonds during the formation of the protein-ligand complex. Thus, to choose the most promising compounds between several ligands with similar $\Delta G^0$ the enthalpic term can be used as an indicator which of the compounds forms the strongest H-bonds. In following design cycles the entropic portion can be improved by simply increasing the size of the ligand. Water molecules will be displaced and released to the bulk which increases the degrees of freedom of this system. However, these principles have to be used reluctantly since the loss in degrees of freedom of rotatable bonds of the ligand might influence the entropic portion unfavorably. A similar unfavorable influence on $\Delta G^0$ can be assumed if polar groups of the protein residues are shielded by hydrophobic parts of a ligand which is accompanied by a loss in the enthalpic portion.[4, 6] Additionally, the capturing of water molecules is followed by a changed thermodynamic profile. The formed H-bonds between water molecule and ligand usually lead to a gain in enthalpy while the entropy develops unfavorably since the water molecules are fixed in a certain position.[7, 8]

The association and dissociation constants can be measured by surface plasmon resonance (SPR). The structure of the biosensor of this instrument is shown in Figure 1-2.

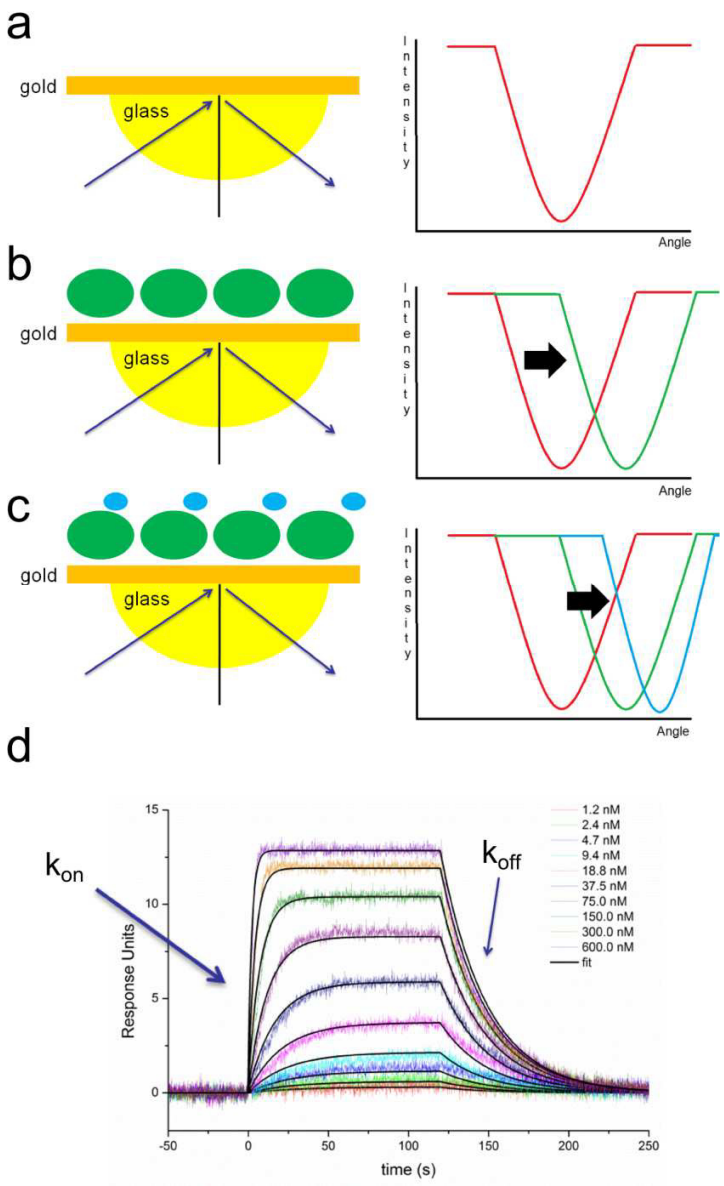



**Figure 1-2 Theory of surface plasmon resonance.** The figure was adapted from www.sprpages.nl.[9] a) Under the SPR angle a dip in the reflected beam of light is detected due to the generation of surface plasmons in the gold layer. b) Green circles represent protein molecules. The refractive index changes due to changes in the buffer next to the gold layer, e. g. due to the attachment of protein molecules. Consequently, the SPR angle changes (green dip). c) Ligand molecules are depicted as blue circles. The binding of ligand molecules to the protein leads to another change in the SPR angle. d) Increasing concentrations of a ligand are measured. The black curves are fitted by the evaluation software to the experimental data.

A thin layer of gold is in tight contact to a glass prism on the one side and to a flow cell which is flushed by buffer on the other side. A beam of light is directed on the glass prism under the angle of total internal reflection. In the gold layer the light interacts with free oscillating electrons (plasmons). Due to this phenomenon a dip in the intensity of the reflected beam of light is detected under the SPR angle (right side of Figure 1-2). The SPR angle is dependent on the refractive index in the buffer close to the surface of the gold layer. Different procedures are available to attach protein molecules to the gold layer. In the amine coupling procedure a mixture of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and N-hydroxysuccinimide (NHS) are used to activate the carboxymethyl dextran surface which generates reactive succinimide esters. Afterwards a solution of the protein is injected over the surface of the protein. The amine groups of the lysine residues of the protein molecules react with the succinimide esters which attaches them covalently to the sensor chip. This changes the SPR angle (green dip in Figure 1-2b). Remaining succinimide esters are deactivated with ethanolamine. For the measurement of kinetic constants increasing concentrations of the inhibitors are injected over the sensor chip surface for a certain time. Afterwards buffer flows over the sensor chip so that the ligand can dissociate from the protein. The binding of the ligand is visible by the change of the SPR angle (blue dip in Figure 1-1c). This change is plotted (Response Units, on the y-axis) against the time (on the x-axis), resulting in a SPR sensorgram. With an evaluation software a suitable binding model can be fitted onto the measured sensorgrams (Figure 1-2d) which gives the kinetic constants $k_{on}$ and $k_{off}$. The $K_D$ value is obtained by the ratio of $k_{off}/k_{on}$.[10, 11]

The half-life of a receptor-ligand complex is defined as $0.693/k_{off}$. This half-life is directly related to the residence time of the ligand complexed to the protein. Since inhibitors which exhibit a similar $K_D$ value can show different values for $k_{off}$ and $k_{on}$ the kinetic constant can be used to classify the screening hits. Dependent on the intended pharmacological use a short or a long residence time is desired. For example, for drugs which lower the blood pressure a slow dissociation (long residence time) is advantageous. On the contrary, drugs which could have toxic side effects should have a short residence time.[10, 12]

## 1.3 Human Aldose Reductase: a Target for Drug Design to Treat Diabetic Complications and a Model System for Structural, Thermodynamic and Kinetic studies

### 1.3.1 Introductory remarks about diabetes mellitus

The World Health Organization defines diabetes mellitus as a hyperglycemic condition with consequent microvascular complications.[13] Symptoms for hyperglycemia are e. g. polyphagia, polydipsia, polyuria, fatigue and weight loss.[14] Diagnostic criteria for diabetes mellitus from the American Diabetes Association (ADA) are A1C ( $HbA_{1C}$, glycated hemoglobin) $\geq$ 6.5%, fasting plasma glucose (FPG) $\geq$ 126 mg/dl, 2-h plasma glucose (2-h PG) $\geq$ 200 mg/dl or classical syndroms of hyperglycemia or hyperglycemic crisis together with a random plasma glucose level $\geq$ 200 mg/dl.[15] Four different types of diabetes mellitus can be differentiated. Type I relies on a complete malfunction of the β-cells of the pancreas. In type II a progressive insulin secretory deficiency develops due to an insulin resistance of the cells. Type III is the gestational diabetes which can break out in the second or third trimester of a pregnancy. In type IV several forms of diabetes due to other underlying causes are summarized. The former paradigm which classified the onset of type I diabetes in young age and the one of type II in older age seems not to be justified anymore since recent studies indicate that both types are present in both age groups.[15, 16]

### 1.3.2 Prevalence of diabetes

It is estimated that in 2013 382 million people worldwide suffered from diabetes. This means that 8.3% of all adults suffered from this disease. Diabetes is no longer the disease of only the industrial countries since 80% of the cases are found in the low- to middle-income countries, making this disease a big threat for the health of the world population. Estimations predict an increase up to 592 million people living with diagnosed diabetes in 2035.[17] Additionally, it implies a big significance for the health care systems which have to attend to such a huge amount of patients with a chronic disease.

### 1.3.3 Treatment of diabetes

Type I diabetes has to be treated with insulin. The use of multiple dose injections which consists of basal and prandial insulins is recommended which should mimic physiological insulin secretion. The time of injection should be well coordinated with the carbohydrate intake, the blood level of glucose and the intended physical activity.[18] For the treatment of type II diabetes metformin is the preferred pharmacological ingredient. If the single therapy with metformin is not effective enough a second and also a third substance could be added to the therapy. Therefore numerous further oral antidiabetic agents are available: Sulfonylureas, thiazolidinediones, DPP-4 inhibitors, SGLT2 inhibitors and GLP-1 receptor agonists. In some cases the additional application of insulin is indicated for type II diabetes as well.[18, 19]

### 1.3.4 Diabetic complications

Patients who suffer from diabetes have a high risk of developing complications which are related to the underlying sickness. Cardiovascular diseases like angina, myocardial infarction, stroke, peripheral artery diseases and congestive heart failure might occur. Diabetes is one of the major causes of chronic kidney diseases and of blindness. Small blood vessels are damaged and thus the supply of the organs is disturbed. Peripheral neuropathies develop in the extremities, especially the feet.[20-23] As a consequence pain, tingling and loss of feeling develop. Especially the latter constitutes a major problem in diabetic patients. Injuries are not noticed which can lead to serious infections and ulcerations. This can result in the development of the diabetic foot disease and lower limb amputations. Peripheral nerve damage develops in 25% of patients 10 years after diabetes is diagnosed and in 50% of patients after 20 years. This diabetic peripheral neuropathy influences walking speed, stability, balance and reaction time, negatively.[22, 24] Since 30- 50% of type II diabetes remains undiagnosed the diabetic complications must be considered as even more meaningful because these patients have no possibility for preventive actions.[25, 26] A recent review classified diabetes as the epidemic of the century and highlighted the need for extensive research especially to prevent the development of chronic complications.[23, 27]

### 1.3.5 Pathobiology of diabetic complications

#### 1.3.5.1.1 Role of the human aldose reductase in the development of diabetic complications

Only certain kinds of cells are pathologically modified in patients with diabetes: Capillary endothelial cells in the retina, mesangial cells in the renal glomerulus, neurons and Schwann cells in the peripheral nerves. These cells, in contrast to other cells, cannot control the intake of glucose which leads to high internal glucose levels.[21]

Already in 1966 Gabbay et al. could show the significance of the sorbitol pathway for the development of diabetic neuropathy.[28] It was found that the levels of glucose, sorbitol and fructose are elevated in peripheral nerves and spinal cord in mildly diabetic animals. Cell membranes are relatively impermeable for sorbitol and fructose. The authors concluded that these substances are produced from glucose using the sorbitol pathway (Figure 1-3).

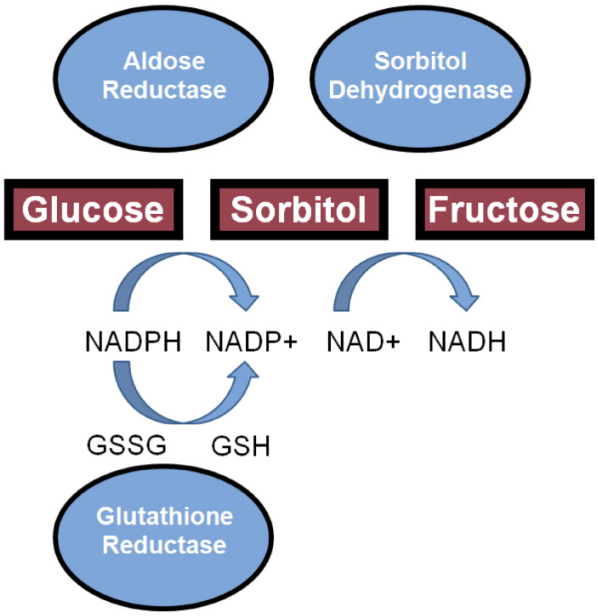



**Figure 1-3 The sorbitol pathway.** The human aldose reductase catalyzes the first step of the reaction. NADPH is needed as a cofactor. Adapted from Brownlee.[20]

In the first step, human aldose reductase (ALR2, EC 1.1.1.21) reduces glucose to sorbitol using the cofactor NADPH. In the second step of the pathway, sorbitol dehydrogenase oxidizes sorbitol to fructose using $NAD^+$ as a cofactor. ALR2 reduces toxic aldehydes to inactive alcohols. However, in a diabetic patient the concentration of glucose in the cell becomes too high and ALR2 uses it as a substrate. Previously it was assumed that the elevated sorbitol levels would lead to a high osmotic pressure. But sorbitol does not reach such high concentrations. The depletion of NADPH is now regarded as the important pathological consequence of the sorbitol pathway. NADPH is additionally the cofactor for the regeneration of reduced glutathione. The latter is an important intracellular antioxidant. Thus, the activity of the ALR2 leads to a higher vulnerability for intracellular oxidative stress.[21] Engerman et al. measured the effect of the aldose reductase inhibitor sorbinil on diabetic dogs over a treatment period of five years.[29] The administration of sorbinil prevented the decline of the nerve conduction velocity which was observed in not treated diabetic dogs. This shows that ALR2 plays an important role in the development of diabetic complications.[21]

Today we know that development of diabetic complications is a complex process. At least three additional mechanisms accompany the polyol pathway: increased formation of advanced glycation end products (AGEs), hyperglycemia-induced activation of protein kinase C (PKC) isoforms and increased hexosamine pathway flux.[21] This might explain why the sole administration of an ALR2 inhibitor only shows moderate effects in clinical studies.[30, 31] Consequently, the combination of an ALR2 inhibitor with an inhibitor of one of the other three pathways is a promising therapeutic option.[32]

### 1.3.6 Human aldose reductase as a target for drug design

The ALR2 consists of 315 residues and contains a single peptide chain. It has a β/α-barrel motif which is built up by a central core of eight parallel β-sheets (yellow in Figure 1-4) which are connected by eight α-helices (red in Figure 1-4) running anti-parallel to the β-sheets.[33] The resulting TIM-barrel is an unique structural motif for an oxidoreductase since most reductases have a Rossman fold. The substrate binding site lies at the C-terminal end of the TIM-barrel which is bordered by several of the loops which connect the α-helices to the β-sheets. $NADP^+$ is stabilized across the barrel. Its nicotinamide is located at the bottom of the binding site. The

binding and release of the cofactor is regulated by the safety belt loop (shown in blue in Figure 1-4) which is formed by the residues 213-226.[34, 35]

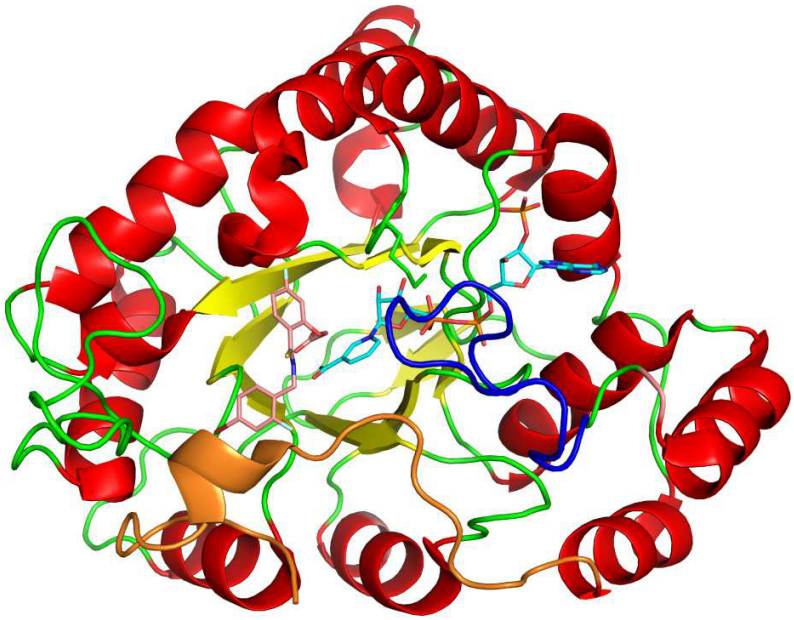


**Figure 1-4 Schematic Representation of ALR2 complexed with an inhibitor and $NADP^+$ (PDB: 1US0).** α-strands are shown in red, β- sheets in yellow and loops in green. The safety belt loop (residues 213-Ser226) is shown in blue and the terminal loop (residues 291-315) which implies the residues which are important for the flexibility of the specificity pocket, in orange. Carbon atoms of the inhibitor are shown in pink and of $NADP^+$ in cyan. The oxygen atoms of the inhibitor and $NADP^+$ are shown in red, the nitrogen atoms in blue. This figure and the following representations of protein structures in this work were generated using PyMOL.[36]

This loop can translocate by 17 Å (in relation to the position of the α-carbon of Trp 219) to allow for the entry of NADPH or the exit of $NADP^+$.[35] The C-terminal loop of the ALR2 (orange in Figure 1-4) borders the bottom side of the binding pocket. The binding site of ALR2 can be separated into rigid and flexible areas (Figure 1-5).[37] The part which is very conserved among the structures of various ALR2 ligand complexes consists mainly of Tyr48, His110, Trp111 and the nicotinamide moiety of $NADP^+$. It is referred to as the anion binding pocket because it can

stabilize the negative charge of an inhibitor or a substrate with the help of the positively-charged cofactor and different H-bonds to the aforementioned residues.[38]

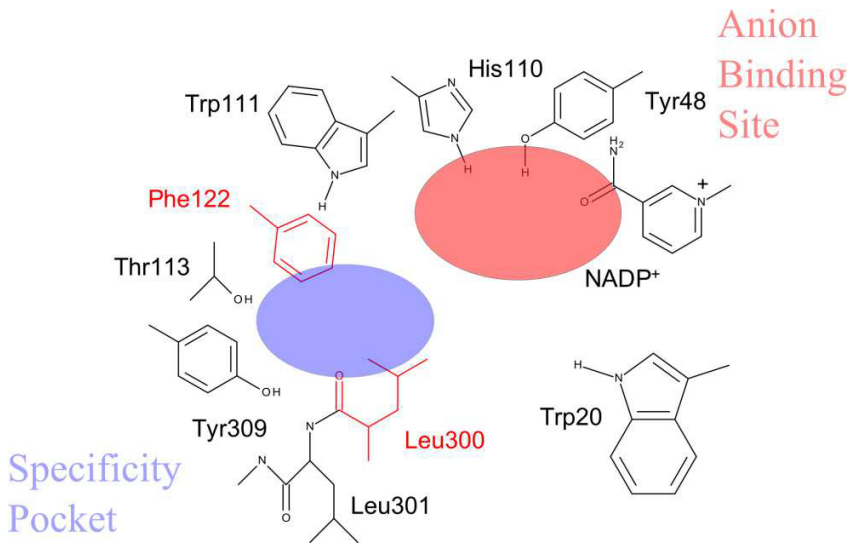



**Figure 1-5 Schematic view on the binding pocket of ALR2.** The anion binding pocket is highlighted in red and the specificity pocket in blue. Phe122 and Leu300 (shown in red) serve as gatekeepers and can open the entrance to the specificity pocket.

On the contrary the residues which are bordering the specificity pocket exhibit a high flexibility. This is best visible for Leu300 and the connected residues. The side chain of Leu300 serves together with Phe122 as gatekeepers to open and close the entrance to the specificity pocket.[37] For the design of ALR2 inhibitors the addressing of this pocket appears promising. In aldehyde reductase, which is closely related to ALR2, the opening of a respective pocket is hindered since it would involve the rupture of a salt bridge. Consequently, ligands with moieties which are binding in the specificity pocket are selective for ALR2.[39]

### 1.3.7 Human aldose reductase as a model system

The search for a potent inhibitor made ALR2 a well characterized protein. The published complex structure with the inhibitor IDD594 (0.66 Å, PDB: 1US0)[40] is still one of the PDB entries with the highest resolution. A protocol for overexpression in *E. coli* is available. Several thermodynamic studies were performed.[8, 41] Details about the protonation changes could be elucidated.[42] Inhibitor molecules bearing different chemical structures, especially different head groups which address the specificity pocket, are described in literature.[43-45] Consequently, this

protein is an excellent system to study the background of the process of bimolecular recognition between the protein and ligands.

## 1.4 Human Carbonic Anhydrase II as a Model System for Biophysical Studies

Carbonic anhydrases (CA) are present, as far as we know, in all organisms.[46] Four families of carbonic anhydrases (CA) exist (α-, β-, γ- and δ-CAs). In mammals, 16 isoforms of α-CAs or carbonic anhydrase related proteins (CARP) have been discovered.[47] The human carbonic anhydrase II (hCAII, EC 4.2.1.1.) belongs to the α-CAs and is a cytosolic enzyme. It functions as a zinc hydrolase which is catalyzing the simple reaction of $CO_2$ and water to bicarbonate and a proton:[48]

$$H_20 + C0_2 \rightleftharpoons H^+ + HCO_3^-$$

This reaction is reversible and depends on the environmental pH value. hCAII is with a turn-over rate of $10^6$ $s^{-1}$one of the fastest working proteins.[48, 49] Due to this simple reaction the protein is involved in several basic physiological processes, e.g. pH and $CO_2$ homeostasis, respiration, transport of bicarbonate between different tissues and secretion of electrolytes.[47] hCAII consists of one domain with 259 amino acids. The tertiary structure is dominated by a ten-stranded β-sheet which spans through the protein, dividing it into two halves. The beta sheets have an antiparallel orientatation with the exception of two sheets which are parallel. They are surrounded by several α-helices (Figure 1-6).[50] The binding pocket is conically shaped. At the bottom a $Zn^{2+}$ ion is coordinated by three histidine residues (His94, His96 and His119) and a water molecule which is most likely present as a hydroxide ion.[47] This $Zn^{2+}$ ion catalyzes the reaction by ordering a hydroxide ion in a suitable position to undergo a nucleophilic attack on the $CO_2$ molecule. The formed bicarbonate is subsequent displaced by a water molecule. To start a new reaction cycle this water molecule has to be deprotonated (Figure 1-7a).[47] The fast transport of $H^+$ is assumed to be conducted by a proton shuttle formed by His64, His4, His3; His10; His15 and His17.[47, 51, 52] The binding pocket is amphipihilic. Tyr7, Asn62, His64, Asn67, Gln92, Glu106, Thr199 and Thr200 are forming the hydrophilic area while Val121, Val143, Leu198 and Trp209 are hydrophobic residues (Figure 1-7b).[46, 48]

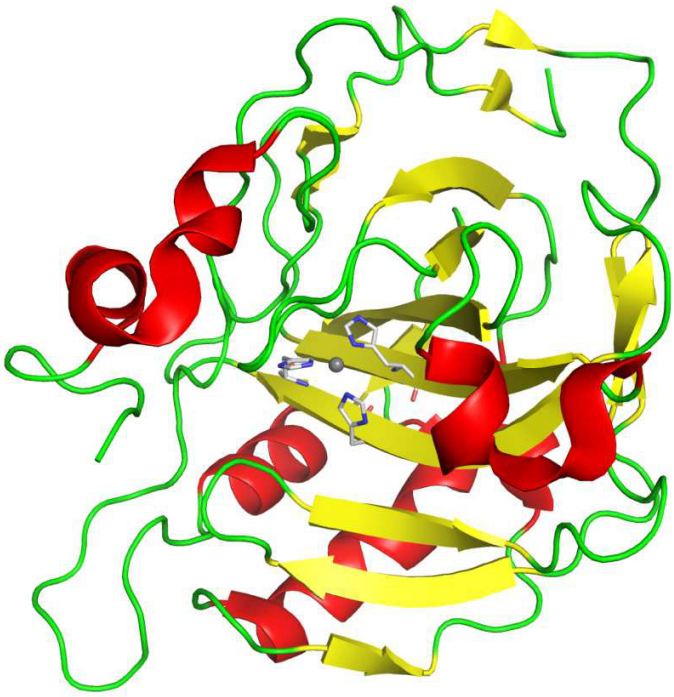


**Figure 1-6 Structure of hCAII (PDB: 3KS3[53]):** α-helixes are shown in red, β-sheets in yellow and loops in green. The residues His94, His96 and His119 are shown as a stick model with carbon atoms colored in grey, nitrogen atoms in blue and oxygen atoms in red. $Zn^{2+}$ is shown as a grey sphere. This figure and the following representations of protein structures in this work were generated using PyMOL.[36]

Initially inhibitors of hCAII were used as diuretic agents. Today, for this indication more potent drugs like high-ceiling diuretics and thiazides are available.[54] Moreover, hCAII inhibitors are established drugs to treat glaucoma, due to their ability to lower the intraocular pressure.[47, 54-56] Furthermore they are used to prevent and treat high altitude disease.[57] Since studies indicate that CA, particularly some of the isoforms beyond hCAII, might be involved in tumor growth the treatment of cancer with hCAII inhibitors might evolve as a future field of therapy.[54] Similar as ALR2 also hCAII can be used as a model system for structural and thermodynamic characterizations of protein-ligand interactions. Highly resolved crystals structures are accessible and ITC data are available for several ligand series.[46, 58, 59] SPR measurements for the bovine form of this protein were described[60, 61] which facilitated the establishment of a protocol for hCAII.

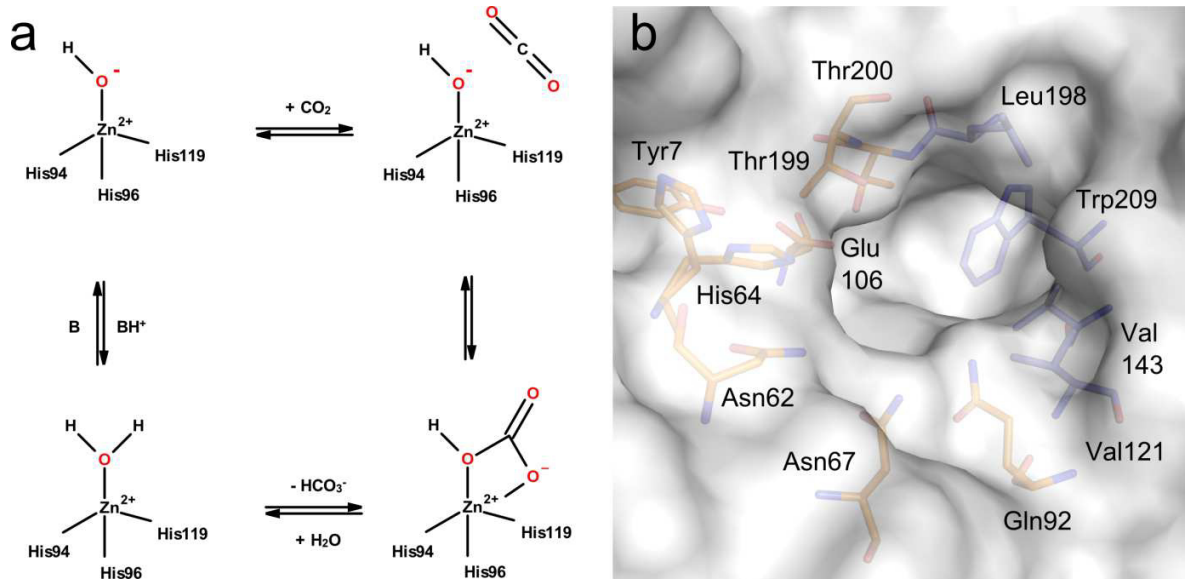



**Figure 1-7 Reaction mechanism and view on the amphiphilic binding pocket of hCAII.** a) Reaction cycle of the reaction of $CO_2$ with water which is catalyzed by hCAII. This figure was adapted from Supuran and Scozzafava.[47] b) Binding pocket of hCAII (PDB: 3KS3[53]). The surface of the protein is colored in grey. Oxygen atoms are colored in red and nitrogen atoms in blue. The carbon atoms of hydrophobic residues are colored in blue and of hydrophilic residues in orange.

## 1.5 Aims of this Work

For the development of drugs it would be preferable to design compounds in a tailor fashion with particular and well-tuned thermodynamic and kinetic behavior. However, the knowledge about these properties and the parameters determining them is still limited. Molecular docking programs still have problems to predict the affinity and even more enthalpy and entropy of a bound ligand to its target protein correctly.[62, 63] A profound understanding about all underlying processes would be necessary to improve the scoring functions of such computer programs. To gain deeper insights, the binding of congeneric ligand series to their respective target proteins can be investigated. The chemical structures of the inhibitor molecules should differ only slightly which allows the assignment of detected changes in kinetics or thermodynamics to a certain structural variation. Based on such data further insights into the process of molecular recognition between proteins and ligands can be obtained which might be employed for similar drug design projects.

For this reason, several projects have been carried out in this work:

**The structural and thermodynamic characterization of an ALR2 inhibitor with a new scaffold:** ALR2 is a promising target to prevent the development of diabetic complications. Epalrestat is the only ALR2 inhibitor which is currently marketed (Kinedak®, Eparel®) in Japan and India.[64] A new type of inhibitor based on a carboxymethylated mercaptotriazinoindole (CMTI) scaffold showing high selectivity compared to aldehyde reductase and AKR1B10 was investigated in this work by protein crystallography. ITC measurements were used to determine the stoichiometry of the binding reaction.

**Determination of the thermodynamic profiles of two new compounds with affinity to ALR2 and AKR1B10:** ITC measurements of two inhibitors for ALR2 and a related protein, AKR1B10, which is presently discussed as a putative anticancer target, have been performed. The findings were correlated with the binding modes which were determined previously by X-ray crystallography.

**Carboxylic acids containing two or three aromatic systems as ALR2 inhibitors:** A further design cycle to optimize an ALR2 inhibitor series originating from a virtual screening[65] was performed by the group of Prof. Martin Schlitzer (Philipps University of Marburg). The crystal structures and thermodynamic profiles of these ligands were characterized using protein crystallography, $IC_{50}$ value determination and ITC measurements. The influences of subtle differences in the binding modes on affinity and thermodynamic profiles have been elucidated.

**Investigating the specificity pocket of ALR2:** A congeneric series of ligands based on the chemical structure of the previously published inhibitor IDD393 was used to induce different conformations of the specificity pocket of ALR2. The influence of different substituents on the aromatic ring system which addresses the specificity pocket was investigated. The thermodynamic parameters were determined and set in relationship to the conformation of the specificity pocket.

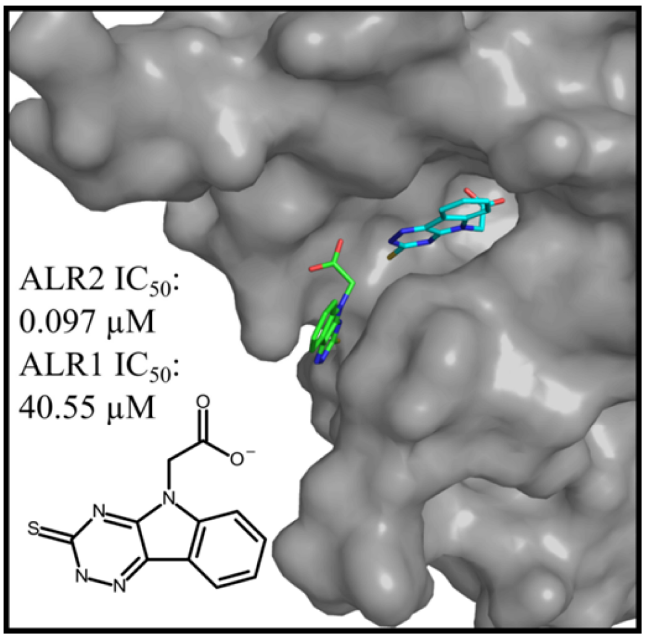
**SPR measurements to elucidate the mechanism of binding of sulfonamides to hCAII:** The kinetic constants and the complex crystal structures of a series of five benzenesulfonamides

binding to hCAII were determined with the aim to propose a possible mechanism of the association process of this ligand series with hCAII. The results were compared with the findings from molecular dynamics (MD) simulations and with previously published data from stopped-flow fluorescence (SFF) experiments.[66, 67]

# 2 Identification of Novel Aldose Reductase Inhibitors Based on Carboxymethylated Mercaptotriazinoindole Scaffold

## 2.1 Introductory Remarks

The following chapter was published in the *Journal of Medicinal Chemistry*.[A] The study has been done in cooperation with Milan Stefek, Marta Soltesova Prnova and Magdalena Majekova from the Slovak Academy of Sciences in Bratislava, Slovakia. The complex crystal structure of compound **13** with ALR2 and the characterization of its thermodynamic profile have been contributed by the author of this thesis along with a significant contribution to the drafting of the manuscript.




[A] Reprinted (adapted) with permission from Stefek, M., Soltesova Prnova, M., Majekova, M., Rechlin, C., Heine, A. and Klebe, G. (2015) Identification of Novel Aldose Reductase Inhibitors Based on Carboxymethylated Mercaptotriazinoindole Scaffold, *J. Med. Chem. 58*, 2649-2657. Copyright (2015) American Chemical Society.

Identification of Novel Aldose Reductase Inhibitors Based on
Carboxymethylated Mercaptotriazinoindole Scaffold

## 2.2 Abstract

Fifteen compounds, sharing an indole-1-acetic acid moiety as a common fragment, were selected from commercial databases for testing aldose reductase inhibition. 3-Mercapto-5H-1,2,4-triazino[5,6-b]indole-5-acetic acid (**13**) was the most promising inhibitor, with an $IC_{50}$ in submicromolar range and high selectivity, relative to aldehyde reductase. Crystal structure of aldose reductase complexed with **13** revealed an interaction pattern explaining its high affinity. Physicochemical parameters underline the excellent „lead-likeness“ of **13**, as a promising candidate for further structure optimizations.

## 2.3 Introduction

In diabetic patients, under conditions of hyperglycemia, glucose becomes a substrate of aldose reductase (ALR2), the first enzyme of the polyol pathway. Increased flux of glucose through the polyol pathway results in overproduction of the organic osmolyte sorbitol. Depletion of NADPH cell stores by aldose reductase may eventually increase the susceptibility of cells to damage by reactive oxygen species. In addition, the polyol pathway contributes to the glycation process supplying fructose, a reactive glycation agent. Through these mechanisms glucose exerts its toxicity with diabetic complications as a consequence.[68, 69]

Aldose reductase, due to its physiological role, metabolizes lipid peroxidation products, contributing in some settings to the inflammatory response. Inhibition of aldose reductase was found to prevent significantly the inflammatory signals induced by cytokines, growth factors, endotoxins, high glucose, allergens and auto-immune reactions in cellular as well as animal models.[70, 71]

The binding pocket of aldose reductase consists of a rigid anion binding pocket and a flexible specificity pocket (Figure 2-1a). The anion binding pocket is formed mostly by His110, Tyr48, Trp111 and the positively charged $NADP^+$. Different classes of inhibitors are described in literature which exhibit varying head groups such as carboxylic acids, succinimides, hydantoins and phenolic OH groups which bind with their negative charge to this pocket.[45] This specificity pocket may only be opened by inhibitors that offer the right interactions. Phe122 and Leu300

serve as gatekeepers and can seal the pocket. The main flexibility is shown by the Ala299-Leu300 backbone and is best visible in different conformations of the Leu300 side chain which determines whether the pocket is open or not.[37] In Figure 2-1b the protein is shown complexed with citrate (PDB: 1X96)[34] with the specificity pocket in closed conformation. In Figure 2-1c, structure of ALR2 complexed with the highly potent aldose reductase inhibitor 2-[(4-Bromo-2-fluorobenzyl)carbamothioyl]-5-fluorophenoxy acetic acid (**20**,[40] Figure 2-1d), is shown with the specificity pocket in open conformation (PDB: 1US0).

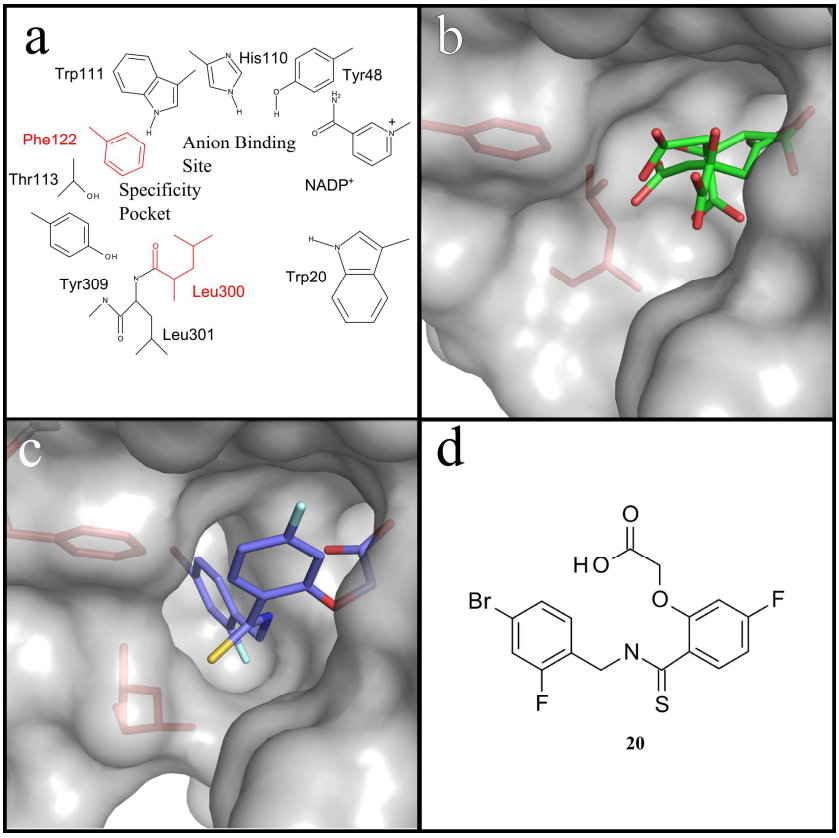
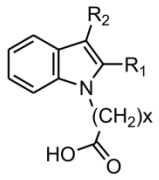



**Figure 2-1**(a) Schematic view on the binding site of ALR2. The anion binding pocket is mainly formed by Tyr48, His110, Trp111 and $NADP^+$. The specificity pocket is opened and closed because of the conformational changes in the backbone of the Ala299-Leu300 stretch. (b) ALR2 complexed with citrate (PDB code: 1X96). The surface of the protein is shown in gray. Phe122 and Leu300 are colored in red. The carbon atoms of citrate are colored in green and the oxygen atoms in red. Two conformations of citrate are visible. The specificity pocket is observed in closed state. (c) ALR2 complexed with **20** (PDB code: 1US0). The carbon atoms of the inhibitor are shown in blue and the oxygen atoms in red. The protein surface is colored in gray. Phe122 and Leu300 are colored in red. The side chain of Leu300 is in a conformation that opens the specificity pocket. The ligand binds presumably with a deprotonated carboxylic acid function to the anion binding pocket. (d) Chemical structure of **20**.

Identification of Novel Aldose Reductase Inhibitors Based on Carboxymethylated Mercaptotriazinoindole Scaffold

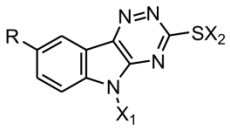
Aldose reductase inhibitors (ARIs), such as acetic acid derivatives (tolrestat, zenarestat), spiro hydantoins (sorbinil), or the succinimide class of compounds (ranirestat) have been primarily investigated for their role against diabetic complications.[72-74] None of them has demonstrated sufficient efficacy in human clinical trials without undesirable side effects. The involvement of

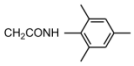
a



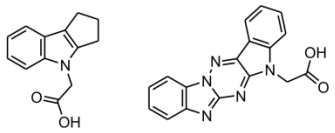
| Compound | x | $R_1$ | $R_2$ |
|---|---|---|---|
| **1** | 1 | H | H |
| **2** | 2 | H | H |
| **3** | 1 | H | $COC(CH_3)_3$ |
| **4** | 1 | H | $CH_2N(CH_2CH_3)_2$ |
| **5** | 2 | H | $CH_2N(CH_2CH_3)_2$ |
| **6** | 1 | $CH_3$ | $CH_2N(CH_3)_2$ |
| **7** | 1 | H |  |
| **8** | 1 | H |  |

c

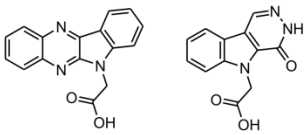


| Compound | R | $X_1$ | $X_2$ |
|---|---|---|---|
| **13** | H | $CH_2COOH$ | H |
| **14** | H | $CH_2COOH$ | $CH(CH_3)_2$ |
| **15** | H | $CH_2COOH$ |  |
| **16** | H | $CH_2COOH$ |  |
| **17** | $CH_3$ | $CH_2COOH$ | $CH_2CONHCH_2CH_2CH_3$ |
| **18** | H | H | $CH_2COOH$ |
| **19** | H | H | H |

b



9 10



11 12

**Figure 2-2 (a) Compounds 1- 8. (b) Compounds 9 - 12. (c) Compounds 13 - 19.**

ALR2 in inflammatory pathologies apparently renewed the interest in the field of ARIs.[75] In the search for better ARIs, the focus has been shifted towards novel chemotypes. Emphasis was given not only to the inhibitory activities but also to the physicochemical properties determining biological availability in target tissues.[76, 77] Recently presented examples of a new generation of aldose reductase "differential" inhibitors were found to preferentially inhibit the reduction of either hydrophilic or hydrophobic substrates.[78, 79] Substituted indole-1-acetic acids represent a group of ARIs of high activity and selectivity,[80-83] with lidorestat as a promising candidate for clinical studies. Yet lidorestat was withdrawn from these studies owing to its side effects. Nonetheless, we still consider the indole-1-acetic acid moiety as a promising starting fragment for the design of efficient, selective and pharmacologically applicable ARIs.

In line with the above findings, we used a ligand-based drug design strategy to search for novel ARIs in databases of commercially available compounds. Fifteen compounds with an indole-1-acetic acid moiety as a common fragment in their structure were selected for experimental testing of their inhibitory potency against aldose reductase. Among the compounds studied, 3-mercapto-5H-1,2,4-triazino[5,6-b]indole-5-acetic acid (compound **13**) exhibited the highest biological activities both at the level of isolated enzyme and at organ level, in a model of isolated rat eye lenses. The stoichiometry of the binding reaction was determined by ITC. To identify crucial interactions within the aldose reductase binding site, the crystal structure of ALR2 in complex with **13** was determined. Selectivity in relation to the closely related rat kidney aldehyde reductase (ALR1) was evaluated. Structure–activity relationships along with a computer-based physicochemical profiling of the compounds are discussed.

## 2.4 Results and Discussion

### 2.4.1 Compounds

ChemSpider database was screened for compounds exhibiting the structural requirements to embed an indole-1-acetic and indoline-1-acetic acid moiety. An initial database of 5813 compounds was thus obtained. Coincidence of four additional requirements, namely experimental/ACD calculated logP = 0-3.5, free N-carboxymethyl group, up to 6 hydrogen bond donors and 2-6 hydrogen bond acceptors reduced the set to 1147 compounds. Then applying

LASSO similarity value of 0.99 for ALR2 inhibition in combination with commercial availability requirement gave 335 compounds.

In our further search we focused on indole-1-acetic acids bearing a neutral substituent (compound **3**) or a basic substituent (compounds **4-8**) in position 3 (structural analogues of lidorestat). In addition, we were interested in the effect of embedding the substituent into a fused ring (compounds **9-17**). Vendor accessibility was taken into consideration in selection of the compounds for experimental evaluation. Their structures are shown in Figure 2-2. For comparison, two indole-1-propionic acid congeners (**2, 5**) were included into the experimental sample set. In addition, two structural variants lacking the indole-1-acetic acid moiety were selected (compounds **18, 19**).

### 2.4.2 ALR2 inhibition

The compounds **1- 19** were evaluated for their potential to inhibit the *in vitro* reduction of D,L-glyceraldehyde by partially purified ALR2 extracted from rat eye lenses, using zopolrestat and epalrestat as reference inhibitors (Table 2-1). It has been shown that human ALR2 exhibits 85% sequence homology to rat lens ALR2,[84] while the catalytic active sites of both enzymes are considered identical.[85] Unsubstituted indole-1-acetic acid (**1**) inhibited ALR2 in low micromolar range as suggested by the measured $IC_{50}$ values. Insertion of a substituent at position 3 (compounds **3,4, 6-8)** resulted in reduced inhibitory power towards ALR2 compared to the unsubstituted indole-1-acetic acid (**1**)**.** Variation of the basic substituents in compounds **4**, **6-8** had only mild effect on the inhibitory potential. For the propionic acid congeners **2** and **5**, a profound decrease of inhibition was observed, with $IC_{50}$ ~50 μM for **2** and $IC_{50}$>100 μM for **5**. The polycyclic derivatives **9** and **10,** with one or two additional fused aromatic rings, respectively, revealed moderately increased inhibitory power compared to **1**. Yet, the addition of a third ring in **11** resulted in a substantial drop of inhibitory potential. Increased polarity of the fused ring in **12** resulted in a similar reduction of the inhibitory power. On the other hand, profoundly increased inhibition was recorded for the mercapto triazine derivatives **13** and **17**, with $IC_{50}$ values in submicromolar range.

**Table 2-1 Inhibitory Effect of Compounds 1-19 on rat ALR2 and ALR1.**

| | | $IC_{50}$ (μM)[A] | |
|---|---|---|---|
| **compound** | **MW** | **ALR2** | **ALR1** |
| **1** | 175.19 | 7.30 ± 0.03 | 80.16 ± 2.30 |
| **2** | 189.22 | 53.54 ± 3.70 | 11.75 ± 0.96 |
| **3** | 259.31 | 13.76 ±0.73 | 68.44 ± 3.48 |
| **4** | 260.34 | 30.14 ± 0.67 | 38.05 ±2.13[C] |
| **5** | 274.37 | 19.85 ± 0.14[C] | 6.65[C] |
| **6** | 246.31 | 42.16 ± 2.06 | 22.31 ± 1.93[C] |
| **7** | 258.32 | 46.48 ± 11.88 | 19.20 ± 0.27[C] |
| **8** | 274.32 | 34.52 ± 1.29 | 32.49 ± 1.26[C] |
| **9** | 215.25 | 2.87 ± 0.86 | 4.82 ± 0.40 |
| **10** | 277.28 | 2.75 ± 0.78 | 24.79 ± 0.61 |
| **11** | 317.3 | 9.95 ± 0.35 | 15.13 ± 2.84[B] |
| **12** | 243.22 | 11.46 ± 2.38 | 48.99 ± 9.04[C] |
| **13** | 260.28 | 0.097 ± 0.019 | 40.55 ± 2.05 |
| **14** | 302.35 | 2.05 ± 0.05 | 14.13 ± 0.36 |
| **15** | 368.39 | 2.24 ± 0.14 | 1.18 ± 4.34[B] |
| **16** | 435.5 | 4.06 ± 0.10 | 23.48 ± 2.11[B] |
| **17** | 373.43 | 0.32 ± 0.05 | 19.56 ± 5.38 |
| **18** | 260.27 | 3.46 ± 0.53[B] | n.d. |
| **19** | 202.24 | 4.41 ± 2.94[C] | n.d. |
| **Zopolrestat** | 419.38 | 0.005 ± 0.001 | 0.049 ± 0.016 |
| **Epalrestat** | 319.4 | 0.25 | n.d. |
| **Sodium - Valproate** | 166.11 | n.d. | 56.1 ± 2.7 |

[A] Results are mean values ± SD from at least three measurements.
[B] %I at 10 μM concentration.
[C] %I at 100 μM concentration.

Values of $IC_{50}$ in low micromolar range were recorded for the remaining mercapto triazine derivatives (**14-16**). The strongest inhibition of ALR2 was found for **13** with $IC_{50}$ ~ 0.1 μM. In addition, we obtained a closely related value ($IC_{50}$ = 0.057 ± 0.002 μM) for the inhibition of human recombinant AKR1B1 by **13**, measured under identical reaction conditions as applied for rat ALR2 assay, with the protein concentration of 0.28 μM. The structure-activity relationship in the series of **12, 13, 18** and **19** points to the necessity of a concurrent presence of both, the carboxymethyl group in position 5 and the terminal sulfur to achieve best affinity. Shifting of the carboxymethyl group from indole nitrogen to sulfur in **18** or its absence in **19** resulted in a loss of activity.

### 2.4.3 Selectivity

An important feature of pharmacologically applicable ARIs is their high selectivity over other members of the aldo-keto reductase family. The co-inhibition of structurally related physiologically relevant oxido-reductases may create undesired side effects. To test for selectivity, the closely related aldehyde reductase (ALR1) is routinely studied.[86] The ALR1 inhibitory potencies, expressed as $IC_{50}$ values for the inhibition of the reduction of the substrate D-glucuronate by partially purified ALR1 from rat kidney, are shown in Table 2-1. For comparison, the reference valproic acid is included. All tested compounds were found to be less active on ALR1 compared to ALR2. In the series of the most efficient ARIs, the selectivity indices were found to be the highest for the mercapto triazines. It should be noted that the most selective compound **13,** with a selectivity factor above 400**,** is also the most potent ARI in this series.

For compound **13**, AKR1B10 inhibition was assayed yielding an $IC_{50}$ value of 21.4 ± 2.0 μM and the corresponding selectivity factor with respect to ALR2 was above 200.

### 2.4.4 Enzyme kinetics

In the next step, the enzyme kinetics for the most efficient ARI **13** was analyzed. Uncompetitive inhibition was observed in relation to D,L-glyceraldehyde as a substrate (Figure 2-3) with a corresponding inhibition constant $K_i$(**13**) = 0.089 ± 0.018 μmol/L.

### 2.4.5 Conformation analysis

For the most efficient ALR2 inhibitor **13,** tautomeric forms were first estimated by the Marvin calculator[25] in the range of pH = 6.2 – 8.0. As a result, three tautomers were selected (Figure 2-4) and investigated by means of quantum-chemical calculations. The most preferred tautomer in solution was found to be **13-1** followed by the forms **13-2** and **13-3** with relative energy stabilities of 16.4 kJ/mol and 72.9 kJ/mol, respectively (see Section 2.6.4).

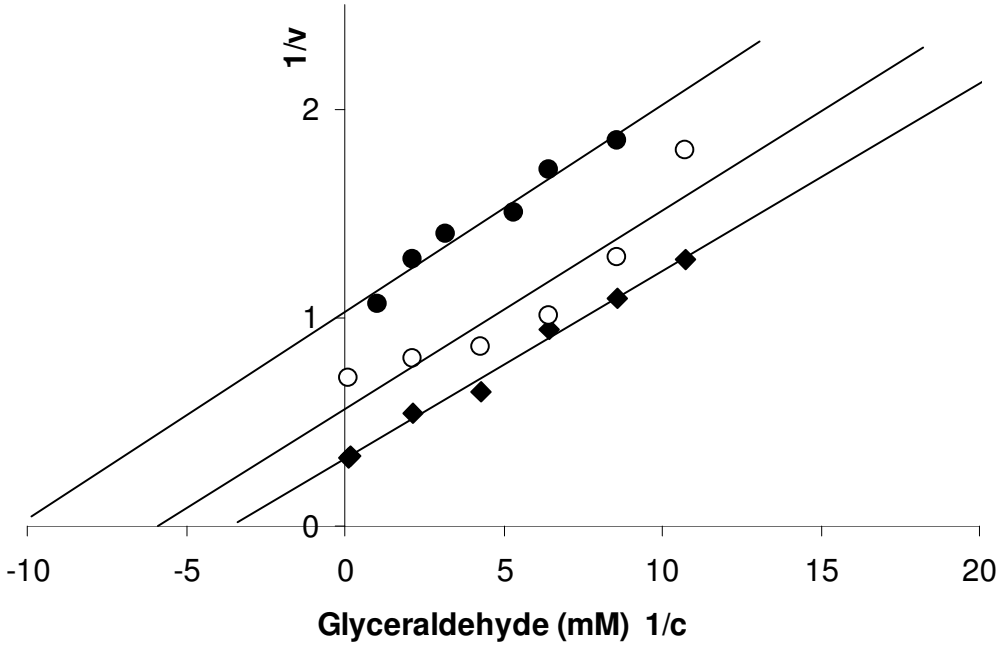



**Figure 2-3 Inhibitory effect of compound 13 on rat lens aldose reductase.** Typical double reciprocal plots of the initial enzyme velocity versus the concentration of substrate (*D,L*-glyceraldehyde) in the presence or absence of **13**: (♦) no inhibitor; (○) 0.05 μM; (●) 0.1 μM of **13** (uncompetitive type of inhibition).

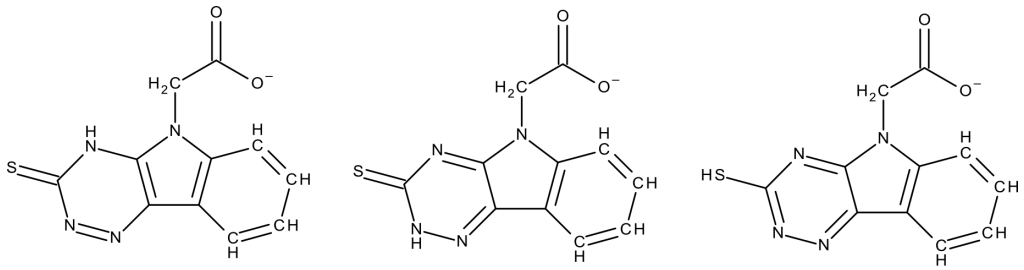



**Figure 2-4 Tautomeric forms of compound 13.**

### 2.4.6 Crystal structure and isothermal titration calorimetry (ITC) measurements

The crystal structure of ALR2 in complex with **13** was determined at pH 8.0 with a resolution of 1.26 Å. Two molecules of the inhibitor are visible in the determined difference electron density, one as expected in the binding pocket and a second in the vicinity of the entrance of the pocket

(Figure 2-5a). The latter is in contact with the binding site inhibitor molecule (first molecule) and refines to an occupancy of 70% whereas the active site molecule shows 100% occupancy. The oxygen atom of the carboxylate group of the second inhibitor molecule and N2 of the first one form an H-bond (2.8 Å, Figure 2-5c). The proton is most likely contributed by the nitrogen of the triazine ring of the first molecule of **13,** as we assume that the carboxylate is deprotonated under the applied conditions.

In our structure (Figure 2-5b), the first inhibitor molecule which occupies the anion binding pocket forms similar interactions as described for related inhibitors exhibiting a carboxylate head group, for example the inhibitor **20** (IDD594), with which the so far best resolved crystal structure of an ALR2-inhibitor complex could be determined.[40] One carboxylate oxygen atom of the inhibitor forms an H-bond to NE2 of His110 (2.7 Å) and to the Tyr48OH group (2.7 Å), while the second oxygen is in contact with NE1 of Trp111 (3.0 Å). The most likely negatively charged carboxylate group of **13** makes an electrostatic interaction with the oxidized and then positively charged $NADP^+$ (3.1 Å). A water molecule is connected to ND1 of His110 (3.0 Å). One further water molecule is interacting with N1 of the triazine ring of the inhibitor (3.0 Å). The remaining contacts to the protein are of hydrophobic nature, e.g. to Trp219 (3.5 Å), Trp20 (3.5 Å), Trp111 (4.0 Å), Trp79 (5.0 Å), Val47 (3.8 Å), Phe122 (4.2 Å). Presumably the inhibitor binds favorably to this aromatic environment due to its own aromatic character.

We analyzed the crystal structure to rationalize the good selectivity of **13** for ALR2 compared to ALR1. The conformation adopted in the binding pocket is surprising. Previously it has been proposed that addressing the protein with the specificity pocket in open state is a prerequisite for a high selectivity index,[39] however we found in the present complex this pocket to be closed (Figure 2-5a). The distance of the sulfur atom of **13** to the backbone nitrogen of Leu300 (3.2 Å) indicates a putative H-bond to this residue (Figure 2-5b) even though the geometry of this interaction is not ideal for an H-bond. This requires the sulfur atom to be in the tautomeric state to act as an H-bond acceptor. To obtain accurate values for the bond lengths and angles with corresponding standard deviations in order to support our assumption about the tautomeric state actually adopted for binding to ALR2, we performed a refinement with SHELXL without imposing any geometry restraints for **13**. However, no conclusive results could be obtained.

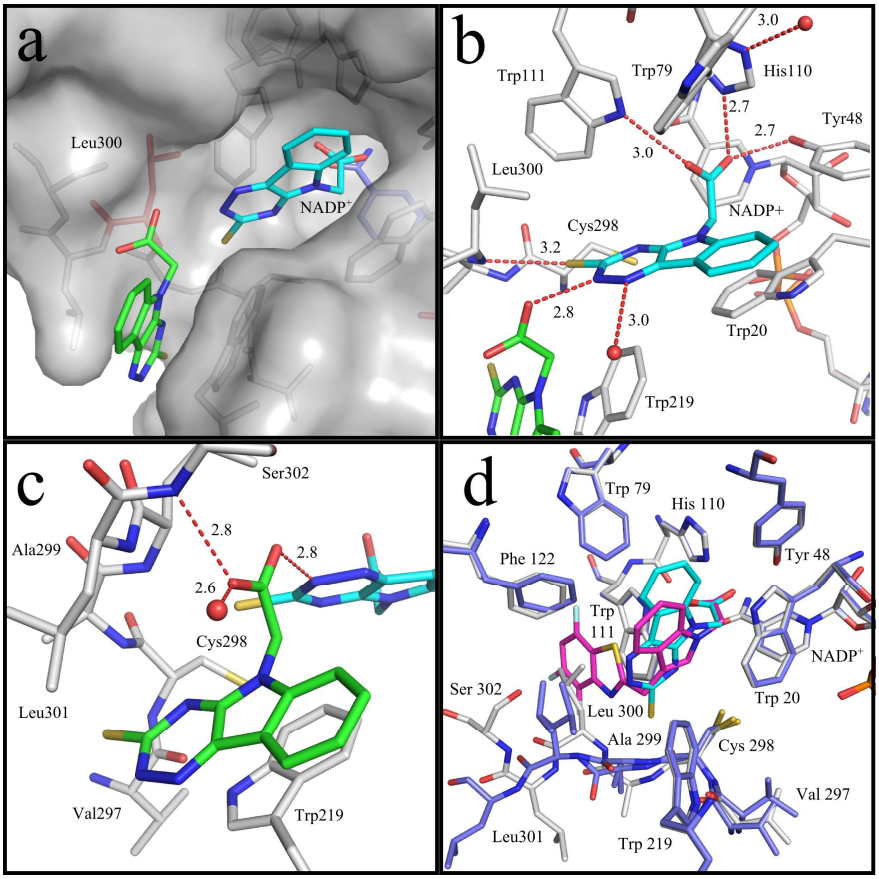



**Figure 2-5** (a) Two molecules of **13** are located in the binding site of ALR2. For $NADP^+$, the first and the second molecule of **13**, the nitrogen atoms are shown in blue, oxygen atoms in red and the sulfur atoms in yellow. For $NADP^+$, carbon atoms are shown in blue, for the first molecule of **13** in cyan and for the second molecule of **13** in green. All atoms of the protein are shown in gray except Leu300 which is colored in red. Leu300 is found in a conformation that closes the specificity pocket. (b) Binding mode of the first molecule of **13**. Nitrogen atoms are shown in dark blue, oxygen atoms in red and the sulfur atoms in yellow. Waters are shown as red spheres. H-bonds are indicated as red dashed lines. Atom color codes are maintained throughout the following representation. The oxygen atoms of the carboxylate group of the inhibitor form H-bonds to Tyr48, His110 and Trp111. The sulfur atom of **13** forms an H-bond to the backbone nitrogen atom of Leu300. The N2 of the triazine ring makes an H-bond to the second molecule of **13** while the N1 is in contact with a water. (c) Binding mode of the second molecule of **13**. The main interactions of the second molecule of **13** are an H-bond between one of the carboxylate oxygen atoms to the N2 of the first molecule of **13**, while the other oxygen atom forms H-bonds to the backbone nitrogen of Ser302 and to a water molecule. (d) Superposition of **13** with lidorestat. The carbon atoms of lidorestat are shown in purple. The carbon atoms belonging to the protein residues of the ALR2–lidorestat complex are shown in blue. Lidorestat opens in contrast to **13** the specificity pocket.

This is most likely due to the insufficient resolution of the dataset to accurately determine the bond length with reasonable estimated standard deviations. Thus, based on the bond lengths, we cannot unambigously assign which tautomer of **13** binds to the protein. Nonetheless, we assume

that tautomer **13-2** is present as it would allow to establish the observed H-bond to the second molecule of **13**. In tautomer **13-1**, N2 cannot serve as an H-bond donor to interact with the most likely deprotonated carboxylate group of the second molecule of **13** (pH 8.0). The quantum-chemical calculations however suggested the tautomer **13-1** to be the most likely one in water. The energies of **13-1** and **13-2** in vacuum were very similar, slightly shifted in favor of **13-2** (Table 2-6 in Section 2.6.4). We even assume the equilibrium between different tautomers might be changed upon ligand binding to the protein as the inhibitor molecule in the anion binding pocket is largely desolvated. The second inhibitor molecule binds (Figure 2-5c) via the described H-bond to the first one and the other carboxylate oxygen atom is in H-bond distance to the backbone nitrogen of Ser302 (2.8 Å). Furthermore, the aromatic core skeleton makes a face to face π-stacking to Trp219 (3.4 Å).

We perfomed ITC measurements to determine the stoichiometry of the binding reaction, which reveals a value in agreement with a 1:1 binding, similarly observed for other inhibitors. Accordingly, the second inhibitor molecule is only weakly bound and possibly results from the applied concentrated crystal soaking conditions. It is well accommodated in the protein structure and likely suggests an additional binding position that could intermediately be adopted during the penetration of the inhibitor into the catalytic center. Also the occupancy of 70% indicates that the second molecule is only weakly bound. In agreement with the first molecule of **13,** we decided to build it as tautomer **13-2**. Our ITC experiments were performed from three buffer conditions (Table 2-2) suggesting a proton entrappment upon protein binding (Figure 2-6).

**Table 2-2 $\Delta H_{obs}$ for 13 in the corresponding buffer system.**

| **buffer** | **$\Delta G^0$** | **$\Delta H^0_{ion}$ [A]** | **$\Delta H^0_{obs}$** |
|---|---|---|---|
| | **[kJ/mol]** | **[kJ/mol]** | **[kJ/mol]** |
| HEPES | 35.9 +/- 0.3 | 21.07 | -51.2 +/- 1.4 |
| Tricine | 34.9 +/- 0.4 | 31.97 | -38.7 +/- 0.6 |
| TRIS | 35.9 +/- 0.2 | 48.07 | -33.1 +/- 0.8 |

[A] Values for $\Delta H_{ion}$ were taken from literature.[87, 88]

This method has been described previously[42, 87, 88] Therefore the enthalpy of binding had to be corrected for buffer dependencies. The thermodynamic parameters are $\Delta G^0$= -35.6 kJ/mol, $\Delta H^0$= -62.8 kJ/mol and $-T\Delta S^0$= +27.2 kJ/mol. The binding process is strongly dominated by the enthalpy and comprises an unfavorable entropic contribution. Solely considering the rigid molecule an enthalpic signature appears surprising. However, the entire binding process including changes of the protein and the residual solvation structure has to be regarded. On the other hand, the strong exothermic signal suggests the formed H-bonds to be quite attractive.

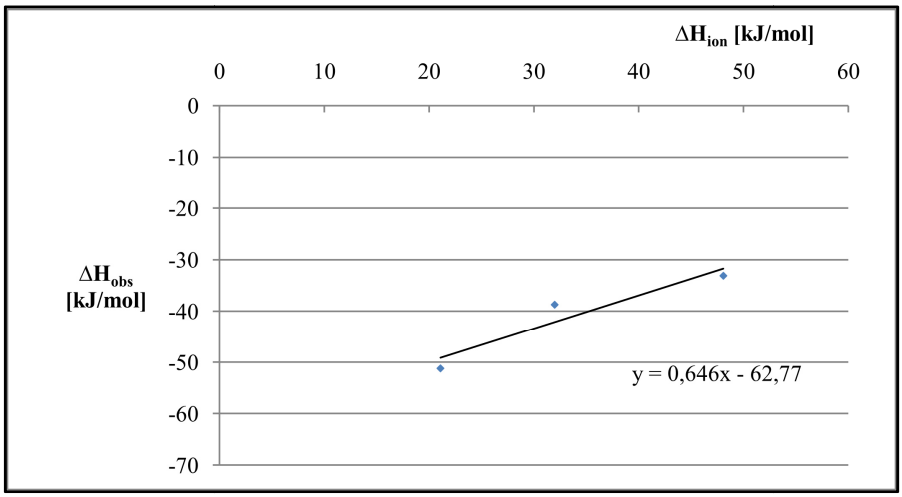



**Figure 2-6 $\Delta H^0_{obs}$ (y, in kJ/mol) plotted against $\Delta H^0_{ion}$ of the applied buffer solution (x, in kJ/mol).** An uptake of 0.65 protons per mole of formed complex during the binding process is indicated by the positive slope.

The contribution resulting from the displacement of water molecules is difficult to estimate in the present case since there is no *holo*-structure of ALR2 available. The uptake of 0.65 protons per mole of formed complex falls into the same range as already described by Steuber et al. for carboxylate-type ALR2 inhibitors.[42] In their study Tyr48 was identified as the proton acceptor.

### 2.4.7 Comparison with the binding mode of lidorestat

The complex crystal structure of lidorestat[80, 81] was aligned with the one of **13** (Figure 2-5d). The carboxylate groups of **13** and lidorestat align perfectly well. Lidorestat opens the specificity pocket with its 4,5,7-trifluoro-1,3-benzothiazole side chain and thus induces the conformational changes in the loop next to Leu300. The aromatic core structure of **13** coincides with the plane as

present in lidorestat. Lidorestat makes several interactions in the specificity pocket which explains the higher $IC_{50}$ value (5 nM) compared to **13**.[21,22]

### 2.4.8 Physicochemical properties

Values of logP and log $D_{7.4}$ calculated by the ACD/Labs software[89] for **13** were found to be 1.68 and -2.03, respectively. The distribution ratio of **13** was also determined experimentally by a standard shake-flask technique which gave a closely similar value of log $D_{7.4}$ =-1.90 ± 0.03.

**Table 2-3 Molecular obesity indices calculated for mercapto-triazino-indoles 13-17.**

| Compound number | $pIC_{50}$ | logP[A] ACD/Labs | logD[A] ACD/Labs | LE[B] | LLE[C] | LELP[D] | BEI[E] |
|---|---|---|---|---|---|---|---|
| **13** | 7.01 | 1.68 | -2.03 | 0.55 | 5.33 | 3.08 | 26.94 |
| **14** | 5.69 | 2.68 | -1.05 | 0.38 | 3.01 | 7.07 | 18.81 |
| **15** | 5.66 | 3.63 | -0.11 | 0.31 | 2.03 | 11.93 | 15.36 |
| **16** | 5.39 | 2.06 | -1.68 | 0.24 | 3.34 | 8.44 | 12.38 |
| **17** | 6.49 | 2.2 | -1.54 | 0.35 | 4.29 | 6.3 | 17.38 |

[A] Log P and log D were calculated with ACD/Labs software[24],
[B] LE = -1.4log($IC_{50}$)/N, N:number of heavy atoms,
[C] LLE = pIC50 – logP,
[D] LELP = logP/LE,
[E] BEI = pIC50/MW (kDa)

Considering the high inhibition efficacy/selectivity along with the structural parameters fitting „the rule of three“, (MW<300 Da, clogP<3), compound **13** appears as a promising lead. Therefore, molecular obesity indices, considering potency, size and lipophilicity in the same index,[90-92] were calculated for **13** and for related mercapto-triazino-indoles **14-17** (Table 2-3). Indeed the calculated values of the corresponding parameters (LE = 0.55, LLE = 5.33, LELP = 3.08 and BEI = 26.94) point to an excellent „lead-likeness“ of **13**, rendering this compound suitable for further optimization. Another aspect that enhances „lead-likeness“ of **13** is its good water solubility (10 mM water solution of **13** could easily be prepared under neutral pH conditions).

### 2.4.9 Eye lens sorbitol

The inhibitory effect of compound **13** with respect to sorbitol accumulation in isolated rat eye lenses was determined. As shown in Table 2-4, increased sorbitol levels were recorded in the isolated rat eye lenses incubated with glucose, in comparison with control incubations without glucose, reflecting increased flux of glucose through the cytosolic lens ALR2. Similarly other authors[93] observed a more than 10-fold increase of sorbitol levels in the isolated eye lenses incubated with glucose under comparable conditions (50 mM glucose, 4 h incubation). Sorbitol accumulation was significantly and in a concentration-dependent way inhibited by compound **13**, starting at a concentration as low as 10 µM. This finding indicates the ready uptake of compound **13** by the eye lens tissue followed by inhibition of the cytosolic ALR2. Obviously, the final response observed at the organ level of isolated lenses is the result of an interplay of well-suited physicochemical properties of **13** (namely lipophilicity and ionization) with good inhibition efficacy.

**Table 2-4 Effect of compound 13 on sorbitol accumulation in isolated rat eye lenses cultivated with high glucose[A]**

| **Incubation** | **Sorbitol** (nmol/g) | **n** |
|---|---|---|
| **-** Glucose | 233.99 ± 7.80 [B] | 15 |
| + Glucose | 772.90 ± 19.70 | 17 |
| + Glucose **+ 13** (1.0 µM) | 668.66 ± 54.15 | 4 |
| + Glucose **+ 13** (10 µM) | 536.01 ± 25.90 [C] | 7 |
| + Glucose **+ 13** (50 µM) | 478.71 ± 30.99 [C] | 4 |
| + Glucose **+ 13** (100 µM) | 395.06 ± 23.83 [C] | 10 |

[A] Glucose, 50 mM; time of incubation, 3 hours; 37 °C. Results are mean values ± SEM from n independent incubations.
[B] $p < 0.001$ vs. (+)Glucose (Student´s t-test)
[C] $p < 0.001$ vs. (+)Glucose (Student´s t-test)

## 2.5 Conclusions

By applying a ligand-based search for novel ARIs in databases of commercially available compounds, fifteen compounds exhibiting an indole-1-acetic acid moiety as common parent scaffold in their structure were selected to experimentally assess their inhibitory potency against

aldose reductase. 3-Mercapto-5H-1,2,4-triazino[5,6-b]indole-5-acetic acid (**13**) with an $IC_{50}$ in submicromolar range and a selectivity discrimination relative to ALR1 about 400 and to AKR1B10 above 200, was identified as promising scaffold for development of efficient, selective and pharmacologically applicable ARIs. The crystal structure determination performed on **13** revealed an interaction pattern explaining its high affinity. The ALR2 inhibitory potential of **13** was demonstrated at the organ level in isolated rat eye lenses incubated with high glucose. Molecular obesity indices, in addition to the structural features matching „the rule of three", along with good water solubility, point to an excellent „lead-likeness" of compound **13**, as a promising candidate for further structure optimizations.

## 2.6 Description of Experimental Procedures

### 2.6.1 Expression and purification, ITC measurements, crystallization and structure determination

The expression and purification of ALR2 are described in Section 9.2.1, the ITC measurements in Section 9.2.5, the crystallization in Section 9.2.3, the soaking in Section 9.2.3.7 and the structure determination in Section 9.2.4.

### 2.6.2 Chemicals

To control the purity of the tested compounds we used LC-MS technique, Agilent 1100 HPLC-MSD SL Ion Trap mass spectrometer with a multi-mode ion source (ESI, APCI) (Agilent Technologies, USA), mostly under following conditions: Column Zorbax SB-C18 1.8 µm 4.6x15mm, Rapid Resolution cartridge (PN 821975-932), mobile phase A – acetonitrile,B – water (0.1% formic acid); flow rate 3 mL/min; gradient: 0 min – 0% A; 0.01 min – 0% A; 0.5 min – 100% A; 1.00 min –100% A; 1.01 min – 0% A; injection volume 1µl; ionization mode - atmospheric pressure chemical ionization (APCI; Scan range m/z 80-1000). For the sample preparation ~0.7 mg of the compound was commonly dissolved in DMSO. A purity of ≥95% has been established for all compounds studied (compounds **1-19**). Sorbitol dehydrogenase, diaphorase, β-$NAD^+$, resazurin, M-199 medium (M 3769), D,L-glyceraldehyde, sodium D-glucuronate, sodium valproate, NADPH, D-glucose, β-mercaptoethanol, epalrestat and $HClO_4$ were obtained from Sigma-Aldrich (St. Louis, MO, USA). AKR1B10 was from Acris Antibodies, Herford Germany. Zopolrestat (lot # 43668-12-7F) was supplied as a gift sample by

Pfizer (Groton, CT, USA). Diethylaminoethyl cellulose DEAE DE 52 was from Whatman International Ltd. (Maidstone, England). Other chemicals were purchased from local commercial sources and were of analytical grade quality.

### 2.6.3 Animals

Male Wistar rats 8- 9 weeks old, weighing 200- 230 g, were used as organ donors. The animals came from the Breeding Facility of the Institute of Experimental Pharmacology Dobra Voda (Slovak Republic). The study was approved by the Ethics Committee of the Institute and performed in accordance with the Principles of Laboratory Animal Care (NIH publication 83-25, revised 1985) and the Slovak law regulating animal experiments (Decree 289, Part 139, July 9th 2003).

### 2.6.4 Computational methods

Calculator Plugins were used for the prediction of the distribution of the compound **13** tautomers.[86] For the most probable tautomers for the range of pH= 6.2 – 8.0, the low energy conformations were obtained by equilibrium conformer systematic search (MMFF94) and subsequent optimization in DFT B3LYP 6-31G* method, all performed in the program SPARTAN'08 (Wavefunction, Inc., Irvine, CA, 2009). To search for the most probable tautomer of **13**, the effect of solvation was also taken into account and the values of total energies $E_{aq}$[94] of the solvated tautomers were calculated, along with the energies E in vacuum.

**Table 2-5 Energies of tautomers of compound 13 calculated in vacuum (E) and in water ($E_{aq}$).**

| | E (kJ/mol) | $E_{aq}$ (kJ/mol) |
|---|---|---|
| **13-1** | -3127206.7 | -3127489.4 |
| **13-2** | -3127207.1 | -3127473.0 |
| **13-3** | -3127173.3 | -3127416.5 |

### 2.6.5 Preparation of rat ALR2

ALR2 from rat lens was partially purified using a procedure adapted from Hayman and Kinoshita[95] as follows: lenses were quickly removed from rats following euthanasia and homogenized in a glass homogenizer with a teflon pestle in 5 volumes of cold distilled water. The homogenate was centrifuged at 10 000 g at 0-4°C for 20 min. The supernatant was

precipitated with saturated ammonium sulfate at 40%, 50% and then at 75% salt saturation. The supernatant was retained after the first two precipitations. The pellet from the last step, possessing ALR2 activity, was dispersed in 75% ammonium sulfate and stored in smaller aliquots in liquid nitrogen container.

### *2.6.6* Preparation of rat ALR1

ALR1 from rat kidney was partially purified according to the reported procedure of Costantino et al.[96] as follows: kidneys were quickly removed from rats following euthanasia and homogenized in a knife homogenizer followed by processing in a glass homogenizer with a teflon pestle in 3 volumes of 10 mM sodium phosphate buffer, pH 7.2, containing 0.25 M sucrose, 2.0 mM EDTA dipotassium salt and 2.5 mM β-mercaptoethanol. The homogenate was centrifuged at 16,000 g at 0-4°C for 30 min and the supernatant was subjected to ammonium sulfate fractional precipitation at 40%, 50% and 75% salt saturation. The pellet obtained from the last step, possessing ALR1 activity, was redissolved in 10 mM sodium phosphate buffer, pH 7.2, containing 2.0 mM EDTA dipotassium salt and 2.0 mM β-mercaptoethanol to achieve total protein concentration of approximately 20 mg/mL. DEAE DE 52 resin was added to the solution (33 mg/mL) and after gentle mixing for 15 min removed by centrifugation. The supernatant containing ALR1 was then stored in smaller aliquots in liquid nitrogen. No appreciable contamination by ALR2 in ALR1 preparations was detected since no activity in terms of NADPH consumption was observed in the presence of glucose substrate up to 150 mM.

### 2.6.7 ALR1 and ALR2 enzyme assays

ALR1 and ALR2 activities were assayed spectrophotometrically as described before[97] by determining NADPH consumption at 340 nm and were expressed as decrease of the optical density (O.D.)/s/mg protein. To determine ALR2 activity, the reaction mixture contained 4.67 mM D,L-glyceraldehyde as a substrate, 0.11 mM NADPH, 0.067 M phosphate buffer, pH 6.2 and 0.05 mL of the enzyme preparation in a total volume of 1.5 mL. The reference blank contained all the above reagents except the substrate D,L-glyceraldehyde to correct for oxidation of NADPH not associated with reduction of the substrate. The enzyme reaction was initiated by addition of D,L-glyceraldehyde and was monitored for 4 min after an initial period of 1 min at 30 °C. ALR1 activity was assayed analogically using 20 mM D-glucuronate as a substrate in the presence of

0.12 mM NADPH in 0.1 M phosphate buffer pH 7.2 at 37 °C. Enzyme activities were adjusted by diluting the enzyme preparations with distilled water so that 0.05 mL of the preparation gave an average reaction rate for the control sample of 0.020 ± 0.005 absorbance units/min. The effect of compounds on the enzyme activity was determined by including in the reaction mixture each inhibitor at required concentrations dissolved in water or DMSO at 1% final mixture concentration. The inhibitor at the same concentration was included in the reference blank. $IC_{50}$ values (the concentration of the inhibitor required to produce 50% inhibition of the enzyme reaction) were determined both from the least-square analysis of the linear portion of the semi-logarithmic inhibition curves and non-linear regression analysis. Each curve was generated using at least four concentrations of inhibitor causing an inhibition in the range from at least 25 to 75%.

### 2.6.8 AKR1B1 and AKR1B10 enzyme assays

The aforementioned procedure for rat ALR2 assay was employed, with the final protein concentrations of 0.28 μM. In AKR1B10 assay, 0.1 M phosphate buffer pH 7.0 was used instead of 0.067 M phosphate buffer, pH 6.2.

### 2.6.9 Eye lens sorbitol assay

The animals in light ether anesthesia were killed by exsanguinations of the carotid artery and the eye globes were excised. The lenses were quickly dissected and rinsed with saline. The compound **13** dissolved in water was added into the tubes containing freshly dissected eye lenses (1 lens per tube) in M-199 medium at pH 7.4, bubbled at 37 °C with pneumoxid (5% $CO_2$, 95% $O_2$), to the final concentrations as reported, 30 min before adding glucose. The incubation was initiated by adding glucose to the final concentration of 50 mM and then continued at 37 °C with occasional (in about 30-min intervals) bubbling the mixture for approximately 30 s periods with pneumoxid. The incubations were terminated after a 3 h period by cooling the mixtures in an ice bath, followed by washing the lenses three times with ice-cold phosphate buffered saline (1 mL). The short term cultivations were preferred to avoid substantial permeability changes of the eye lenses. The washed lenses were kept deep-frozen for sorbitol determination which was performed as described before.[98] In brief, the frozen lenses were let to melt at the ambient temperature. Then distilled water (0.2 mL/1 lens) was added. The lenses were disrupted by a glass rod. The rod was washed twice with distilled water (0.1 mL) and the suspension was ultra-sounded for 5 min.

Thereafter, ice cold $HClO_4$ (9%, 0.4 mL) was added and mixed thoroughly. The mixture was ultra-sounded for another 5 min and then kept on ice for 30 min to let proteins precipitate. The precipitated protein was spun off (15 min at 3 000 rpm) at 4° C. The supernatant was neutralized with concentrated $K_2CO_3$ (4 M). The neutralized supernatant was used for determination of concentration of sorbitol by modified enzymatic analysis.[99] In brief, sorbitol was oxidized to fructose by sorbitol dehydrogenase (SDH) with concomitant reduction of resazurin by diaphorase to the highly fluorescent resorufin. The final concentrations of the assay solutions were: diaphorase (11.5 U/25 mL triethanolamine buffer), $NAD^+$ (25 mg/25 mL triethanolamine buffer), resazurin (0.025 mL 2 mM resazurin solution in 25 mL of triethanolamine buffer), SDH (15.025 U/1 mL triethanolamine buffer). Reaction mixtures were incubated for 60 min at room temperature with an opaque cover. The sample fluorescence was determined at excitation 544 nm, emission 590 nm. After the appropriate blanks were subtracted from each sample, the amount of sorbitol in nmol per gram of lens wet weight in each sample was determined by comparison with a linear regression of sorbitol standards.

### 2.6.10 Partitioning

The distribution ratios D in 1-octanol/buffer systems, defined by total concentration of a solute in organic phase divided by that in aqueous phase, were measured using the shake-flask technique[100] at room temperature. The organic and aqueous phases were mutually saturated. Compound **13** was dissolved in aqueous buffer solution (0.1 M phosphate buffer pH 7.4) in final concentration of 100 μM; the solutions were shaken with 1-octanol for 3 h. Both aqueous and organic phase volumes were 3 mL. The phases were separated by centrifugation for 1 h. The organic layer was removed with a Pasteur pipette. The concentration of the solute was determined in both phases by UV spectrophotometry.

# 3 Structural Determinants of the Selectivity of 3-Benzyluracil-1-Acetic Acids towards Human Enzymes Aldose Reductase and AKR1B10

## 3.1 Introductory Remarks

The following chapter is an excerpt of a manuscript which was prepared for submission to a scientific journal. This project was a collaboration with Dr. Francesc Xavier Ruiz, Alexandra Cousido-Siah, Dr. André Mitschler and Prof. Dr. Alberto Podjarny (CNRS, Illkirch, France), Dres. Marta Domínguez and Ángel R. de Lera (Universidade de Vigo), Dres. Maria Jesús Martin and Jesús Ángel de la Fuente (Biomar Microbial Technologies S. A., León), Dres. Sergio Porté, Isidro Crespo, Jaume Farrés and Xavier Parés (Universitat Autònoma de Barcelona). The author of this thesis contributed the ITC measurements along with the discussion of the corresponding results to this work.

## 3.2 Abstract

Human enzymes aldose reductase (ALR2) and AKR1B10 have been thoroughly related to diabetes, inflammation disorders and cancer. Here we have identified two new lead compounds, JF0048 (**23**) and JF0049 (**24**)**,** that selectively target these enzymes. Although **23** and **24** share the 3-benzyluracil-1-acetic acid scaffold, they have different substituents in their aryl moieties. Inhibition, thermodynamic and structural characterization of both enzymes reveals that the chloronitrobenzene moiety of **23** can open the ALR2 specificity pocket but not the AKR1B10 cognate. In contrast, the larger atoms at *ortho* and/or *meta* positions of **24** prevent the ALR2 specificity pocket from opening due to steric hindrance and provide a tighter fit to the AKR1B10 inhibitor binding pocket.

## 3.3 Introduction

Human aldose reductase (ALR2 or AKR1B1) and AKR1B10 are enzymes of biomedical interest because of their involvement in diabetes (ALR2) and in cancer (ALR2 and AKR1B10). They belong to the aldo-ketoreductase superfamily (AKR) and are NADPH-dependent enzymes

folding into a highly conserved $(\alpha/\beta)_8$ barrel. They have different substrate specificity and inhibitor selectivity due to differences in the residues of their three external and variable loops.[101]

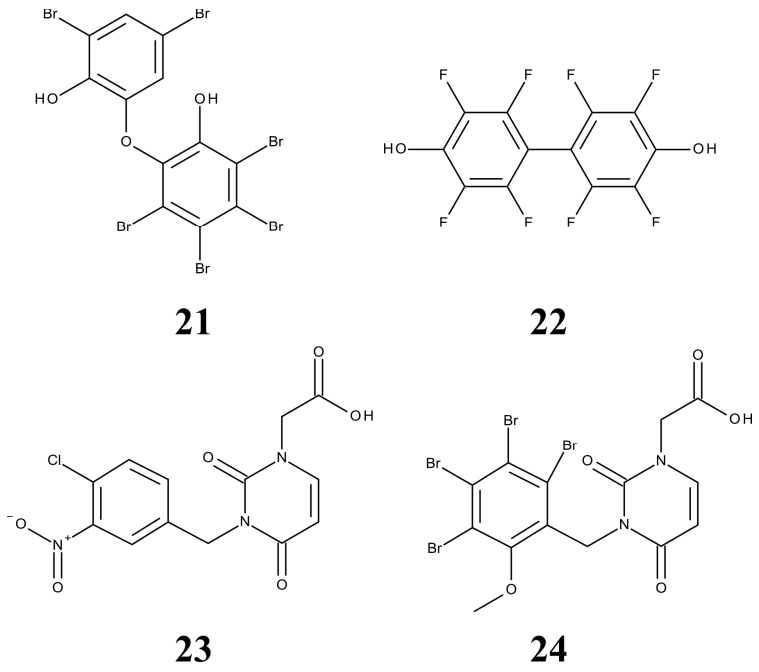



**Figure 3-1 Chemical structures of the investigated inhibitors.**

AKR1B10 shares 71% amino acid sequence identity with ALR2, but its tissue expression is restricted mainly to colon and small intestine. It is overexpressed in lung and liver cancers making it a potential cancer diagnostic and/or prognostic marker.[102-105] AKR1B10 catalyzes the reduction of chemotherapeutic drugs, lipid peroxidation of free aldehydes and all-*trans*-retinaldehyde, but it does not catalyze the reduction of glucose.[106, 107] AKR1B10 may impact the carcinogenesis process through its involvement in several pathways, e.g. retinoic acid signalling,[103, 108-111] lipid aldehyde detoxification and lipid synthesis.[104, 112]

Previously, Biomar Microbial Technologies screened a library of about 2000 marine natural products and found a polybrominated diphenyl ether compound (**21**, Figure 3-1) from a marine sponge (*Disydeaherbacea*) that showed an $IC_{50}$ value of 6.4 µM against ALR2. This hit led to the synthesis of different series of synthetic polyhalogenated compounds (such as **22**), some of which targeted both ALR2 and AKR1B10.[101, 113-115]

In the search of novel ligands selective for AKR1B10, we focused our attention on compound JF0048 (**23**), which was previously shown to be selective for rat lens ALR2 *versus* rat kidney aldehyde reductase.[116] We surmised that the modification of the 3-benzyluracil-1-acetic acid scaffold through the incorporation of additional substituents at the (poly)halogenated aryl moiety, reminiscent of the sponge natural products, might shift the selectivity of the designed compound to target AKR1B10 instead. In the present work, this assumption was indeed confirmed, since compound JF0049 (**24**) (Figure 3-1) was found to be selective for AKR1B10. In addition, structural insights were gained from several X-ray structures of the corresponding complexes. The selectivity of **23** for ALR2 was justified by its ability to open the specificity pocket of the enzyme. On the other hand, the selectivity of **24** for AKR1B10 is probably due to the inability of its bulkier aryl moiety to occupy the specificity pocket of ALR2 and its better fitting into the larger active site of AKR1B10, including a subpocket defined by loop A. This study paves the way for future efforts in structure-guided drug discovery, directed to both ALR2 and the relatively unexplored AKR1B10.

## 3.4 Results and Discussion

### 3.4.1 Inhibition assays of ALR2/AKR1B10 complexes with 23 and 24

Kotani and coworkers[116] had previously assayed **23** against rat lens ALR2 and rat kidney aldehyde reductase, which were purified from the corresponding tissues. They found that **23** was one of the best hits for rat ALR2, with more than 4000-fold selectivity *versus* rat aldehyde reductase. We confirmed that **23** exhibits strong inhibition against human ALR2 ($IC_{50}$ = 170 nM),[117, 118] and is 127-fold and >4000-fold more potent than for AKR1B10 and ALR1, respectively. On the other hand, **24** inhibits AKR1B10 with an $IC_{50}$ value of 450 nM, being 8-fold and >200-fold more potent than for ALR2 and ALR1, respectively (Table 3-1). We have also performed ITC measurements to study the binding of **23** and **24** to both enzymes. As shown in Table 3-3, the Gibbs free binding energy ($\Delta G°$) and the dissociation constants ($K_D$) obtained by ITC are in agreement with the inhibition assays for the binding of **23** and **24** to ALR2. Only for the binding of **24** to AKR1B10, the $K_D$ value is slightly higher in comparison to the corresponding $IC_{50}$ value.

**Table 3-1 $IC_{50}$ values and selectivity ratios.**

| Inhibitor | ALR2 $IC_{50}$ (µM) | AKR1B10 $IC_{50}$ (µM) | ALR1 (100 µM)[A] | AKR1B10 / ALR2 | ALR2 / AKR1B10 |
|---|---|---|---|---|---|
| **23** | 0.17 ± 0.02 | 21.6 ± 1.7 | 5.6% | 127 | 0.008 |
| **24** | 3.6 ± 0.4 | 0.45 ± 0.04 | 27.3% | 0.1 | 8 |

[A] Percentage of inhibition of a 100 µM solution of the respective inhibitor.

### 3.4.2 Crystal Structures of 23 and 24 complexed with ALR2 and 24 complexed with AKR1B10

With the aim to understand the rationale behind the inhibitory properties of **23** and **24** towards ALR2 and AKR1B10, we sought to determine the X-ray structures of the ternary complexes. We were able to obtain all of them, except that of the AKR1B10-**23** complex. The structures fold into the expected $(\alpha/\beta)_8$-TIM barrel, with the cofactor position being conserved.

The X-ray structure of ALR2-$NADP^+$-**23** complex was obtained at the atomic resolution of 1.0 Å, with the $F_o-F_c$ omit map showing clearly the inhibitor electron density (Figure 3-2a). The carboxylic acid of **23** is assumed to be deprotonated, as observed before for **20**,[40, 119] and the oxygen atoms act as hydrogen bond (H-bond) acceptors to Tyr48OH (2.7Å), His110NE2 (2.7 Å) and Trp111NE1 (3.0 Å) and display electrostatic contact with the charged nicotinamide N of $NADP^+$ (4.1 Å).

The uracil-1-acetic acid moiety is performing hydrophobic contacts with Trp20, Trp111, Trp219 and Cys298. Its aryl moiety is sandwiched between the indole moiety of Trp111, through a face-to-face oriented π–π stacking interaction (with a distance of approximately 3.4 Å between the two planes) and the side chains of Phe122, Leu300 and Cys303 lining the specificity pocket. One of the oxygen atoms of the nitro group forms an H-bond with the Leu300 main chain nitrogen atom

(2.9 Å) while the Cl atom undergoes a halogen bond with Thr113OH (2.9 Å, Figure 3-2a). Furthermore, two "non-classical H-bonds" are formed between the oxygen atoms of the nitro group and the Tyr309 $C_{\delta 1}$ and $C_{\varepsilon 1}$ (3.5 Å for both, Figure 3-2a), as observed for other nitro-containing ARIs.[8] Apparently, the sum of the halogen bond and the interactions provided by the nitro substituent at the *meta* position are not triggering a stronger binding. This is not surprising, taking into account that molecular interactions behave in a highly non-additive fashion.[120]

The X-ray structure of ALR2-**24** complex was obtained at 2.45 Å resolution. It has to be noted that crystals were obtained by co-crystallization with 5 mM **24** in a rarely applied reservoir condition for ALR2 crystallization, replacing the habitual buffer ammonium citrate (pH 5.0) with TRIS-HCl (pH 8.0). This is the second ALR2 structure obtained by co-crystallization at this pH value, after the ternary complex of ALR2 with the inhibitor IDD552 (Protein Data Bank (PDB) code 1T41).[121] In addition, the crystal grew in a very unusual space group, I222, only observed before in one ALR2 X-ray crystal structure, a *holo*-enzyme complex (PDB: 1ABN).[122] The $F_o$–$F_c$ omit map shows unequivocally the inhibitor electron density in two different conformations (A and B), where the neighbouring residue Phe122 adopts two different conformations in concordance with the inhibitor positioning (Figure 3-2b and Figure 3-2c). Both conformations of **24** are stabilized by crystallographic contacts with the complex in the neighbouring asymmetric unit. Conformation A of **24** (60% occupancy) is stabilized by the neighbouring conformation A molecule, with a Br –Br contact between the two Br atoms in *para*, while conformation B of **24** is stabilized through halogen bonds with the main chain of the residues of loop A of the complex in the neighbouring asymmetric unit (Figure 3-2d). Therefore, the observed interaction of **24** with the ALR2 *holo*-enzyme confirms that the conformation adopted by **23** in the cognate complex is not possible for the former compound.

In this regard, we have superimposed **24** in the position of **23** in its ALR2 complex, with the methoxy group oriented towards Ala299 or Phe122 (Figure 3-3). In each case, one of the Br atoms at the *meta* position (*ortho* to the methoxy group) would clash against the $CH_2$ group of Trp79 (2.4 Å distance). For the *ortho* substituents, in case the methoxy would be facing Phe122

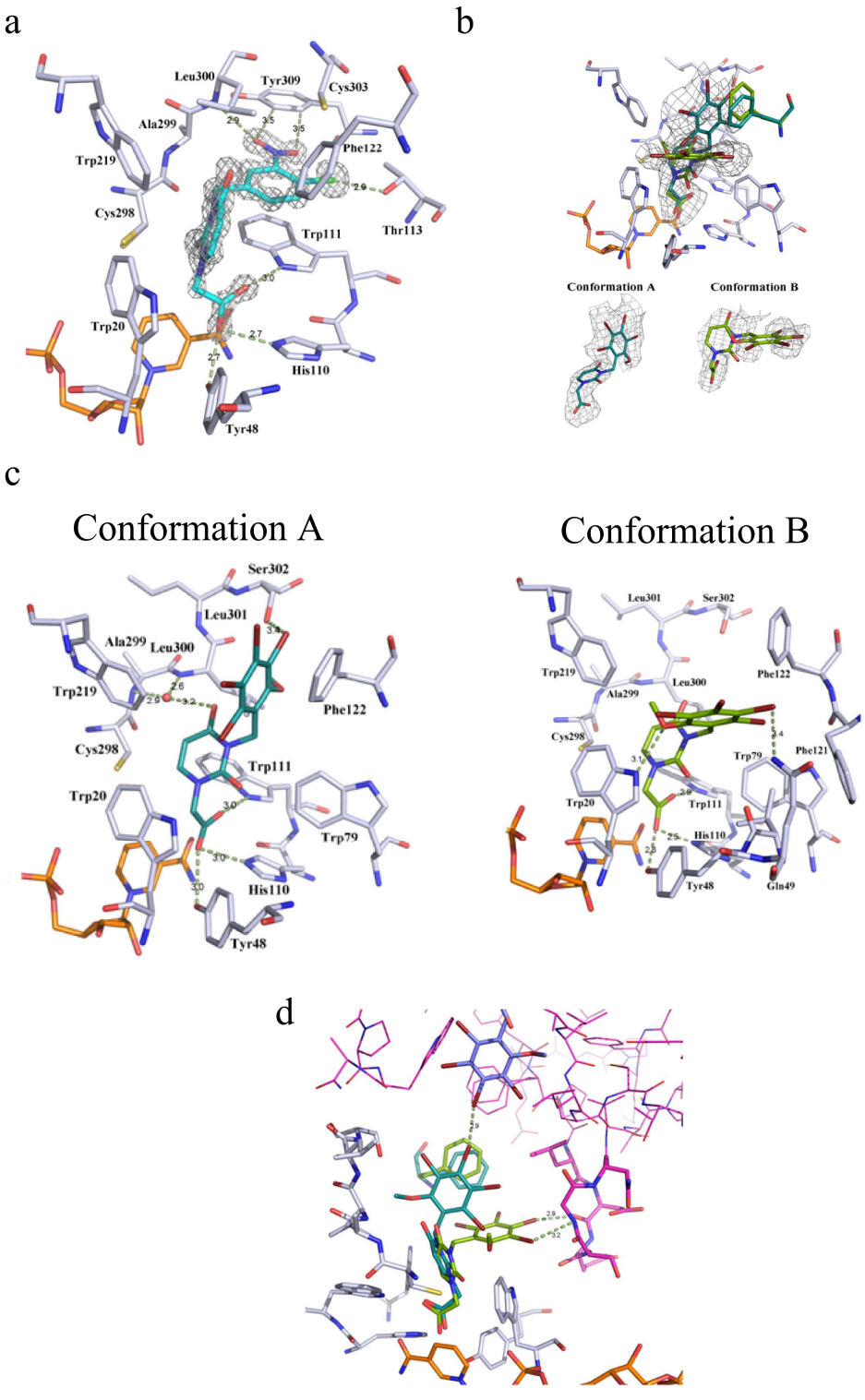



**Figure 3-2** Detailed figure caption on next page.

**Figure 3-2 X-ray structures of ALR2-23 and ALR2-24 complexes.** a) Atomic representation of the inhibitor binding site for the ALR2-**23** complex, with the electron density of the inhibitor shown as a $F_o$-$F_c$ omit map contoured at 6σ level in a grey mesh and the protein residues, the cofactor and **23** in stick representation (white, orange and cyan, respectively); b) Atomic representation of the inhibitor binding site for the ALR2-**24** complex conformation ensemble, with the inhibitor electron density shown as a $F_o$-$F_c$ omit map contoured at 3σ level in a grey mesh (with both conformations showing the inhibitor bound (top) and separated (down)) and the protein residues in white sticks (except for Phe122, which follows the coloring of **24**), the cofactor in orange sticks and **24** in cyan for conformation A and green for conformation B; c) Atomic representation of the inhibitor binding site for the ALR2-**24** complex with each conformation shown in a separate drawing (for color coding see legend to panel b); d) Atomic representation of the inhibitor binding site for the ALR2-**24** complex conformation ensemble including the intermolecular crystallographic contacts with the symmetry-related molecules, in pink lines or sticks for the protein and in blue sticks for the symmetry-related **24**, following the same color code as for the remaining panels. H-bond distances are shown as green-dotted lines.

(Figure 3-3a), a clash would occur with the latter residue (2.1 and 2.2 Å distance with Phe122 $C_{\varepsilon 2}$ and $C_Z$, respectively). If the methoxy group was in the other orientation (Figure 3-3b), there would be clashes with the nitrogen and carbon atoms of the backbone of Ala299 (2.5 Å distance). In the case of AKR1B10, albeit new structures of AKR1B10-inhibitor have been solved in the last two years,[123, 124] we only succeeded to obtain the ternary complex with **24** of the methylated AKR1B10 K125R/V301L mutant (from now on, named as AKME2MU). The same procedure was applied for the determination of the complex structure of inhibitor **22** with AKR1B10.[101, 125] The AKME2MU *holo*-enzyme complexed with **24** was obtained at 1.70 Å resolution, with the $F_o$–$F_c$ omit map showing the inhibitor and the anomalous difference map confirming the position of the four Br atoms of compound **24** (Figure 3-4a). The carboxylic acid oxygen atoms display H-bonds with Tyr49OH (2.9 Å), His111NE2 (2.8 Å) and Trp112NE1 of the conformer A (3.0 Å) (Figure 3-4a). The electron-deficient uracil ring is also involved in further interactions: (i) with Trp21 aromatic moiety through a parallel-displaced stacking (3.5 Å distance) and (ii) the carbonyl oxygen atoms are H-bonded to Phe123 $C_{\varepsilon 1}$ and to Trp220 $CH_2$, the two being "non-classical H-bonds"[8, 126] (Figure 3-4a, atomic distances not shown, being 3.2 Å for both interactions). The methylene bridge and the tetrabromomethoxy aryl moiety are surrounded by a hydrophobic subpocket formed by Trp80, Trp112, Phe116, Phe123, Ala131, Leu301 (Val301 in the wild-type enzyme) and the aliphatic part of Arg125 (Lys125 in the wild-type form) and Gln303 (Figure 3-4b). The most significant interactions are the edge-to-face stacking that the aryl moiety performs with the side chain of Phe123 (3.8Å distance, data not shown) and the interactions of two of the Br atoms of **24** with the aromatic ring systems of Trp80 and Phe116,

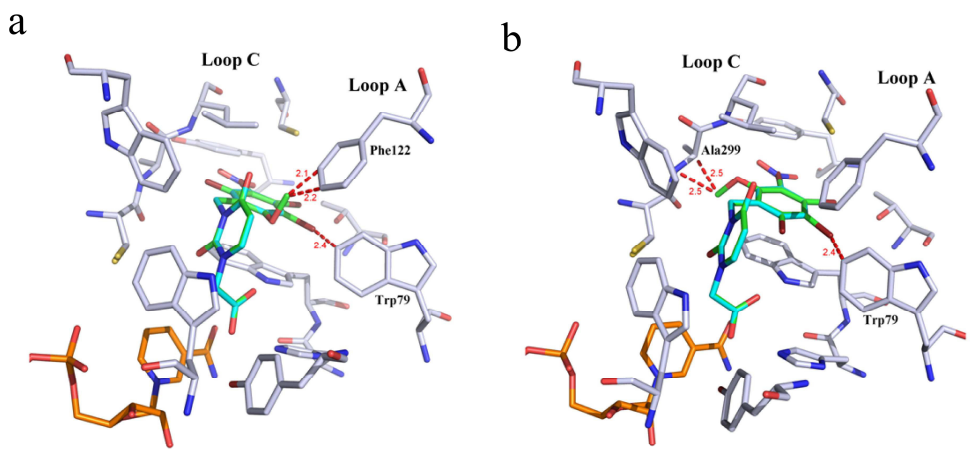



**Figure 3-3 Steric hindrance of 24 within the ALR2-23 active-site conformation.** Atomic representation of the inhibitor binding site of the ALR2-**23** complex with a) **24** superimposed with the methoxy group pointing to loop A, or b) **24** superimposed with the methoxy group pointing to loop C. The protein residues, the cofactor, **23** and **24** are shown as stick representations (white, orange, cyan and green, respectively). The short distances, demonstrating the steric hindrance of **24** in this ALR2 active site conformation, are displayed as red-dotted lines and red labels.

respectively (3.5 Å in both cases, data not shown). Interestingly, Trp112 is observed in two different conformations (Figure 3-4c): conformation A (flipped), with 61% occupancy and positioned as in the complex with the PDB code 1ZUA and conformation B (native), with 39% occupancy and positioned as in the AKR1B10 *holo*-enzyme structure (PDB: 4GQG). Trp112 conformation B is stabilized by a specific H-bond network centred on Gln114 conformation B (58% occupancy) and including Ser304 (Figure 3-4c). The latter conformation had also been observed for several specific AKR1B10 inhibitors.[123, 124] In fact, the occupancy of **24** was found to be 64%, which might suggest that two conformations of the protein were coexisting in the crystal: the conformation A which might be that of the AKME2MU complexed with the inhibitor (with Trp112 flipped), while the conformation B will correspond to that of the unliganded AKME2MU *holo*-enzyme (with Trp112 in a native position). In addition, there is an interstitial

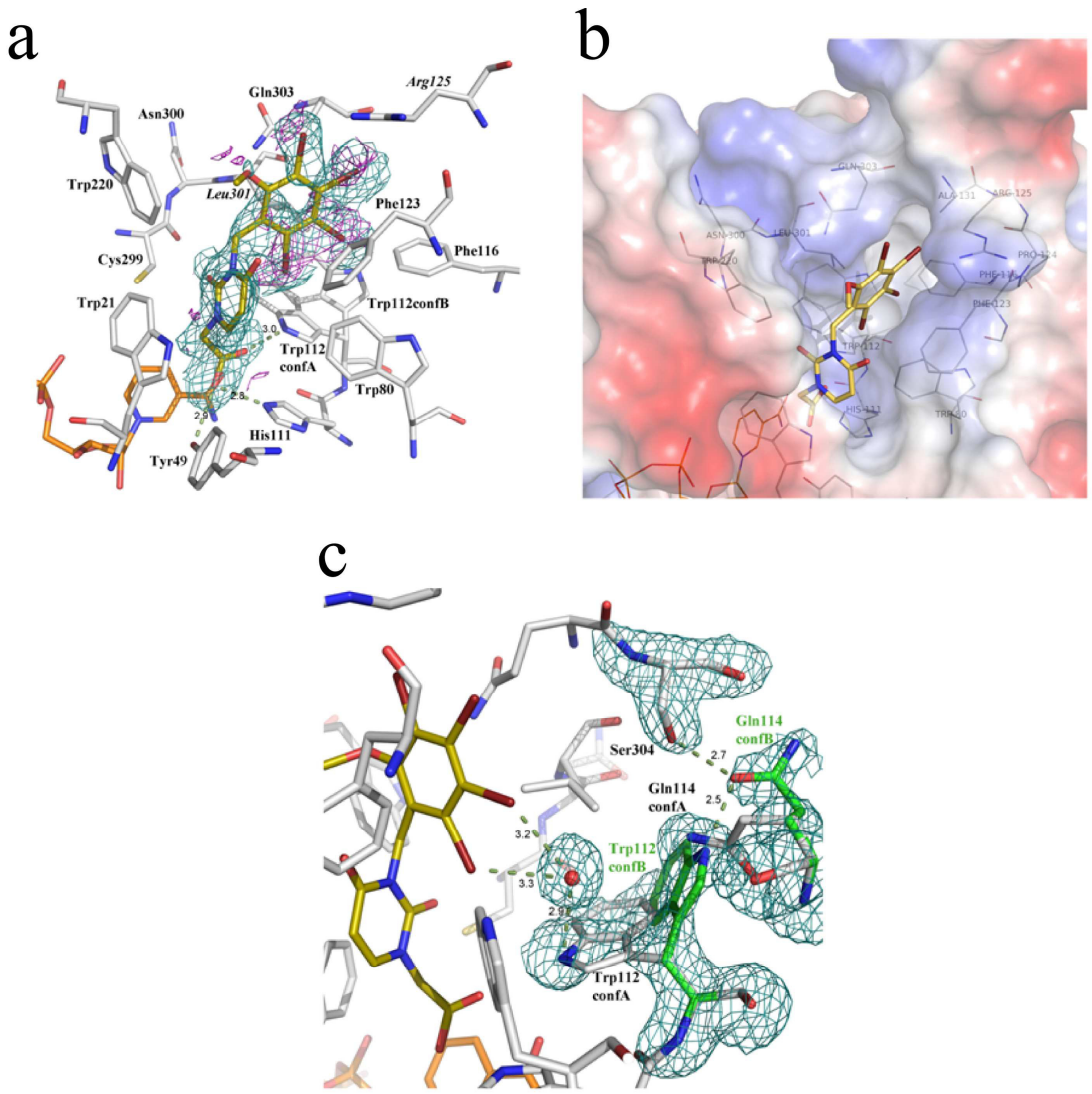



**Figure 3-4 X-ray structure of the AKME2MU-24 complex.** a) Atomic representation of the inhibitor binding site of the AKME2MU-**24** complex, with the inhibitor electron density, shown as a $F_o$-$F_c$ omit map contoured at 2σ level in a cyan mesh and as an anomalous difference map, measured on the bromine edge and contoured at 2σ level in a purple mesh and the protein residues, the cofactor and **24** in white, orange and gold yellow sticks, respectively (mutated residues are labeled in italics); b) Atomic representation of the inhibitor binding site for the AKME2MU-**24** complex, including the surface representation colored according to the local electrostatic potential (blue, positive charge; white, neutral; red, negative charge), calculated with PyMOL, the inhibitor-binding site residues and $NADP^+$ in white and orange lines, respectively and **24** in yellow sticks; c) Detail of the inhibitor binding site for the AKME2MU-**24** complex, following the same color code as in panel a) and including the interstitial water molecule W116 as a red sphere, conformation B of Trp112 and Gln114 in green sticks and the $F_o$-$F_c$ omit map, contoured at 2σ level in a cyan mesh of the aforementioned water molecule, Trp112, Gln114 and Ser304. H-bond distances are shown as green-dotted lines.

water molecule (W116) interacting with two of the Br atoms of **24** (3.3 and 3.2 Å, respectively) and with NE1 (2.9 Å) of the conformation A of Trp122 (Figure 3-4c). This water molecule is also present in the conformation B, as it displays 100% occupancy and does not experience any clashes with the residues of the protein or with the ligand.

### 3.4.3 ITC experiments

ITC experiments were perfomed to gain insights into the thermodynamic profiles of the binding events of the inhibitors to ALR2 and AKR1B10. The binding of the carboxylate-type inhibitors to ALR2 is accompanied by an uptake of 0.7-0.9 protons per mole formed complex by Tyr48.[42, 127] To correct for buffer dependencies, we performed the measurements in three different buffer systems (HEPES, Tricine and TRIS) and plotted the measured enthalpy ($H°_{obs}$) against the heat of ionization ($H°_{ion}$) of the corresponding buffer system. This procedure was previously described in literature.[88] The binding of **23** to ALR2 could be measured in all three buffer systems (Table 3-2). The corrected thermodynamic profile shows a strong favourable enthalpic contribution (-59.7 kJ/mol) while the entropic term is unfavourable (+19.3 kJ/mol, Table 3-3). An entrapment of 0.6 protons per mole is detected (Figure 3-5) which is in good agreement with previous results for similar ARIs. As already mentioned, the formed contacts of the nitro group in the specificity pocket are linked to a strong exothermic enthalpic signal.[8] Presumably, these interactions are responsible for the strong enthalpic binding contribution of **23** to ALR2.

The binding of **24** to ALR2 could only be measured in HEPES since the enthalpic signal was too low to be evaluated in the other two buffer systems (Figure 10-3 in the Appendix). However, a buffer dependency for the binding of **24** to ALR2 could be detected. In HEPES, the binding of **24** to ALR2 is much less driven by the enthalpy than the binding of **23** (-10.7 kJ/mol compared to -46.7 kJ/mol). The measurements of the binding of other carboxylate-type inhibitors to ALR2 which were performed in our laboratory (unpublished results) showed that the enthalpic portion which is corrected for the protonation effect is approximately -11.5 to -14.0 kJ/mol more exothermic than in HEPES buffer. Both types of ligands, those which address and open the specificity pocket and those which do not address it, show this amount of overlaid buffer

dependence in HEPES. It can only be assumed that the binding of **24** is accompanied by a similar proton uptake by the protein. Even if a similar influence on the enthalpic signal due to the proton uptake is assumed as for **23** the enthalpic term of **24** would be much smaller as that of **23**. Thus, the different binding modes of **23** (addresses the specificity pocket) and **24** (leaves the specificity pocket unoccupied) with ALR2 have a significant influence on the thermodynamic profile.

**Table 3-2 $\Delta H_{obs}$ for 23 and 24 in the corresponding buffer system.**

| Buffer | $\Delta H_{ion}$ [A] [kJ/mol] | 23 with ALR2 $\Delta H_{obs}$ [kJ/mol] | 24 with AKR1B10 $\Delta H_{obs}$ [kJ/mol] |
|---|---|---|---|
| HEPES | 21.07 | -46.7 | -38.7 |
| Tricine | 31.97 | -38.6 | -38.2 |
| TRIS | 48.07 | -29.3 | -35.5 |

[A] Values for $\Delta H_{ion}$ were taken from literature.[87, 88]

As **24** binds to ALR2 in two alternative orientations the interpretation of its thermodynamic profile is even more complicated. The following aspects can be seen: in conformation A of **24**, a water molecule is fixed due to H-bonds to the oxygen atom of the benzyluracil-scaffold of **24** (3.2 Å), to the nitrogen atom of the Leu300 backbone (2.6 Å) and to the nitrogen of the Ala299 backbone (2.9 Å). In conformation B, an H-bond between the methoxy substituent of **24** and Trp20 (3.1 Å) can be found. The Br atoms of **24** form contacts to polar groups of ALR2 (Figure 3-2c). In conformation A, such an interaction is visible to the oxygen of Ser302 and, in conformation B, to the nitrogen of Gln49. However, all these interactions do not sum up to a similar contribution to the enthalpy as the (4-chloro-3-nitro-phenyl)-moiety of **23** can account for.

For AKR1B10, only the binding of **24** could be measured since **23** has a too low affinity towards AKR1B10 to be characterized by a direct ITC experiment. Surprisingly, the measurement of **24** with AKR1B10 in the three different buffers revealed that the enthalpic signal of this binding event is much less buffer dependent than it was detected for the binding of **24** to ALR2. A negligible pickup of only 0.1 protons per mole formed complex is detected (Figure 3-5). As described above, Trp112 changes from conformation B to A during the binding of **24** compared

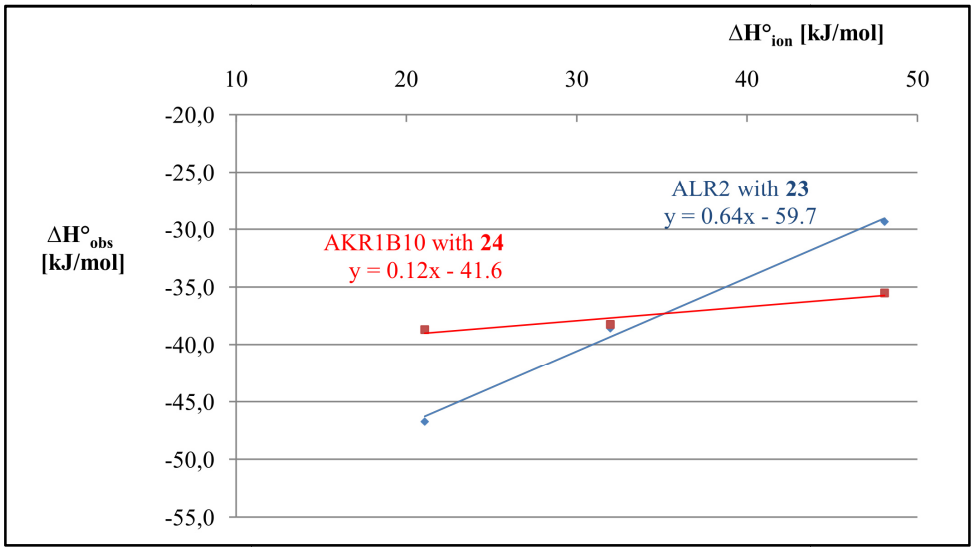



**Figure 3-5 Plot of heat ionization ($\Delta H°_{ion}$) against the measured enthalpy ($\Delta H°_{obs}$) in the corresponding buffer system.** While the binding of **23** to ALR2 shows, as already reported for other ALR2 inhibitors which exhibit a carboxylic acid function, an entrapment of 0.6 protons per mole of formed complex (red line), the binding of **24** to AKR1B10 (blue line) shows an uptake of only 0.1 protons per mole of formed complex.

**Table 3-3 Thermodynamic data of the binding event of 23 and 24 with ALR2 and of 24 with AKR1B10.**

| Complex | $K_D$ (nM) | $\Delta G^0$ (kJ/mol) | $\Delta H^0_{obs}$ [A] (kJ/mol) | $\Delta H^0_{bind}$ [B] (kJ/mol) | $-T\Delta S^0_{obs}$ [A] (kJ/mol) | $-T\Delta S^0_{bind}$ [B] (kJ/mol) |
|---|---|---|---|---|---|---|
| **ALR2-23** | 91.1 ± 39.8 | -40.4 ± 1.3 | -46.7 ± 0.5 | -59.7 | 6.1 ± 0.6 | 19.3 |
| **ALR2-24** | 8894.2 ± 1451.6 | -28.8 ± 0.4 | -10.7 ± 1.9 | n.d. | -18.1 ± 1.5 | n.d. |
| **AKR1B10-23**[C] | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. |
| **AKR1B10-24** | 2313.3 ± 516.7 | - 32.2 ± 0.5 | -38.7 ± 0.7 | - 41.6 | 6.0 ± 0.6 | 9.3 |

n.d.: not determined.
[A] Given is the thermodynamic data for the measurement in HEPES ($\Delta H^0_{obs}$ and $-T\Delta S^0_{obs}$).
[B] If possible, the values were corrected for the protonation effect ($\Delta H^0_{bind}$ and $-T\Delta S^0_{bind}$).
[C] The binding of **23** to AKR1B10 could not be measured due to the low affinity of **23** towards AKR1B10.

to the *holo*-enzyme. This is accompanied by the rupture of the internal H-bond network including Gln114 and Ser304 (Figure 3-4c).[123] A more profound study of the electrostatic properties of the binding pocket of AKR1B10, especially with regard to the structural changes upon ligand binding, is required to quantify the $pK_a$ values of the local residues and their putative changes upon ligand binding to better understand this different behaviour of AKR1B10 compared to ALR2. However, this example underlines the importance of the investigation of buffer dependencies in ITC experiments. The thermodynamic profile of the binding of **24** to AKR1B10, which is corrected for the protonation effect, shows a strong dominance by the enthalpic binding portion (-41.6 kJ/mol) while it is accompanied by an unfavourable entropic term (+9.3 kJ/mol). The enthalpic gain of the newly formed interactions between **24** and the protein residues is obviously overcompensating the enthalpic penalty due to the rupture of the internal H-bond network. As discussed above, we estimate the thermodynamic profile of the binding of **24** to ALR2 to be much less dominated by the enthalpy as is the case for the binding of **24** to AKR1B10 even if the binding profile of **24** to ALR2 would be corrected for the protonation effect. Accordingly, the above described binding mode which **24** adopts with AKR1B10 leads to a more favourable enthalpic contribution. Summarizing both cases, first the comparison of the binding of **23** and **24** to ALR2 and second the comparison of the binding of **24** to ALR2 and AKR1B10, indicate that the opening and occupying of the respective specificity pocket lead to an increase in the enthalpic portion.

## 3.5 Conclusions

A large number of ARIs have been developed for the treatment of diabetic complications. Indeed, huge amounts of preclinical and clinical data have been collected through academic and industrial programs, including a deep structural characterization of the ALR2 binding pocket, which is highly flexible and susceptible to induced fit adaptations. In the last ten years, AKR1B10 has been characterized as a protein related to cell proliferation and shown to be overexpressed in a plethora of different cancers, such as liver, breast, pancreas and lung cancer. ALR2 has also been shown to be involved in processes other than diabetic complications. In fact, it has been established as an important mediator in oxidative stress and inflammation-related processes, including cancer (e.g. colon cancer).[128] Both ALR2 and AKR1B10 siRNA knockdown or drug

inhibition decreased tumor growth of several cell line xenografts (from different tissues) transplanted into immunodeficient mice.[129-133] Therefore, ALR2 and AKR1B10 inhibitors may represent a novel class of antitumor agents and the clinical data assembled in diabetes clinics may help the transition of these inhibitors to cancer chemotherapy.

The AKR1B10 structure has been kept fairly uncharacterized until a couple of years ago and only several new AKR1B10 structures have been deposited in the PDB. Starting from the structure of the uracilacetic acid **23**, we have developed compound **24**, which differs from the former one in the aryl moiety substitution pattern, inspired in this case by the halogenated biaryls and diaryl ethers present in natural products. The inhibition assays, ITC experiments, X-ray crystallographic determinations and *in silico* structural analysis of *holo*-enzyme-**23/24** complexes have allowed dissecting ALR2 and AKR1B10 selectivity and potency.

Several conclusions can be extracted from this work. Starting from the structure of the ARI compound **23** and keeping its uracil-acetic acid unmodified, it has been possible to design and synthesize compound **24**, with an increased volume of the aryl moiety of the N3-benzyl substituent, shifting the selectivity to AKR1B10. The aryl group of **23** is able to open the ALR2 buried specificity pocket located between Trp111 and Leu300, but the bulky substituents at the *ortho* and *meta* positions of **24** prevent its binding into the specificity pocket of ALR2. On the other hand, the large aryl moiety of **24** binds into a more remote hydrophobic subpocket in AKR1B10, located between loops A and C. The nitro group of **23** prevents a proper binding to AKR1B10, as shown by its high $IC_{50}$ value (~20 μM).

The ITC data show that the binding of **24** to AKR1B10 is accompanied by an uptake of 0.1 protons per mole of formed complex, while 0.6 protons per mole are incorporated during the binding of **23** and most likely of **24**, to ALR2, unveiling the difference in electrostatic properties between the two enzyme active sites. The binding of both **23** to ALR2 and **24** to AKR1B10 display a large enthalpic signature while the binding of **24** to ALR2 is more driven by the entropic portion.

## 3.6 Description of Experimental Procedures

### 3.6.1 Site-directed mutagenesis, enzyme expression and purification

The cDNAs of ALR2 and ALR1 were subcloned into the pET15b plasmid, while that of AKR1B10 was subcloned into the pET16b plasmid (pET30-Xa/LIC was used for the K125R/V301Lmutant). ALR2, wild-type and mutant AKR1B10 and ALR1 were recombinantly expressed in the *E. coli* BL21(DE3) strain (Novagen) and were purified using the procedures described previously for ALR2,[134] AKR1B10[108] and ALR1.[135] Purity was confirmed by SDS-PAGE and protein concentration was determined using a NanoDrop ND-1000 (Thermo Scientific).

### 3.6.2 Inhibition screening

The $IC_{50}$ activity assays were carried out according to the quantification of NADPH consumption which takes place when the enzyme catalyses the conversion of D,L-glyceraldehyde into glycerol. The assays were performed at 298 K in 100 mM sodium phosphate buffer, pH 7.0, with protein amount to reach $V_{max}$ and 0.2 mM NADPH. The final reaction volume was 500 μL. All compounds assayed were dissolved in dimethyl sulfoxide (DMSO), with the corresponding solution added to 2% (v/v) DMSO final concentration. They were incubated for 5 min at 25°C prior to addition of the substrate. The reaction was initiated by addition of 1, 60 and 5 mM D,L-glyceraldehyde (for ALR2, AKR1B10 and ALR1, respectively) and the decrease in optical density at 340 nm was monitored for 3 min at 25°C in a UV–Vis spectrophotometer (UV-1700 PharmaSpec, Shimadzu). The $IC_{50}$ value was determined as the compound concentration that inhibits enzymatic activity by 50% and was calculated using the Grafit program (version 5.0; Erithacus Software). Values were given as the mean ± standard error of three experiments.

### 3.6.3 Crystallization and structure determination

Regarding the ALR2-**23** complex, crystals were obtained by co-crystallization under the published condition:[134, 136] hanging drop method at 298 K, reservoir with 50 mM ammonium citrate, pH 5.0, 20% polyethylene glycol (PEG) 6000, with 5 mM of **23**. Neither co-crystallization nor soaking was successful for the ALR2-**24** complex using the previous condition. Later on, co-crystals were obtained using the following condition: 10 mM TRIS-HCl,

pH 8.0, 20% PEG 6000, with 5mM of **24**. Regarding the AKR1B10 complexes, all crystals were obtained with the AKME2MU system (reductive lysine methylation plus K125R/V301L mutations), except for the methylated AKR1B10 *holo*-enzyme. Co-crystals of the AKME2MU-**24** (30 mM inhibitor) complex and of both mutant *holo*-enzymes were obtained by the hanging-drop vapour diffusion method at 298K. The protein solution was mixed with an equal volume of precipitating solution consisting of 100 mM sodium cacodylate, pH 9.0, 30% PEG 6000. A detailed protocol is explained elsewhere.[101] X-ray data were collected on the home source for AKME2MU-**24** complex, at the Swiss Light Source on the X06SA beamline for ALR2-**23** complex and on X06DA for the methylated AKR1B10 *holo*-enzyme and for ALR2-**24**. All data sets were processed with the program HKL2000.[137] The structures were solved by molecular replacement with Phaser[138] and finalized sets of atomic coordinates were obtained after iterative rounds of model modification with the program Coot[139] and refinement with REFMAC5 [140] and PHENIX.[141] Generation of ligand coordinates and restraints were created as previously reported [101] or using eLBOW.[142] The coordinate data for the structures were deposited in the PDB: ALR2-**23**: 4XZH, ALR2-**24**: 4XZI, AKME2MU-**24**: 4XZL, methylated AKR1B10 *holo*-enzyme: 4XZM, AKME2MU *holo*-enzyme: 4XZN. Related figures were prepared with PyMOL (v.1.3; Schrödinger).

### 3.6.4 ITC experiments

The ITC experiments are described in Section 9.2.5. Representative example curves and fitted regression curves are shown in Figure 10-3 in the Appendix.

# 4 Crystallographic and Thermodynamic Characterization of Carboxylic Acids Containing Two or Three Aromatic Systems as Inhibitors of Human Aldose Reductase

## 4.1 Introductory Remarks

The following chapter was prepared for submission to a scientific journal. The synthesis of the compounds was done by Dr. Regina Ortmann and Dr. Michael Eisenmann from the group of Prof. Dr. Martin Schlitzer (Philipps University Marburg). The author of this thesis performed the complex crystal structure determinations, ITC and $IC_{50}$ measurements. Furthermore, he drafted the following manuscript.

## 4.2 Abstract

The human aldose reductase is involved in the development of diabetic complications. Here we present newly synthesized inhibitors bearing a carboxylic acid head group. High-resolved complex crystal structures along with the thermodynamic characterization are presented. Furthermore, we present the first complex crystal structure of an inhibitor of this series possessing an acrylic acid head group and of an inhibitor containing a third aromatic moiety. The findings provide valuable input for a further optimization cycle of this group of inhibitors.

## 4.3 Introduction

382 million people worldwide were afflicted with diabetes in 2013. Up to the year 2035 the estimated number of cases could rise to 592 million.[17] As long-term complications neuropathy, nephropathy and retinopathy can occur.[45] Due to its role in the polyol pathway the human aldose reductase (ALR2, EC 1.1.1.21) is an interesting target for drug design. It catalyzes the reduction of glucose to sorbitol, using NADPH as cofactor. The following depletion of NADPH results in a too low level of reduced glutathione, for which regeneration this cofactor is needed as well. Consequently, the cell experiences a higher risk to suffer from intracellular oxidative stress which leads to the development of diabetic complications.[21] The binding pocket of ALR2 is depicted in Figure 4-1a. The anion binding pocket is formed by Tyr48, His110, Trp111 and the nicotinamide moiety of $NADP^+$.

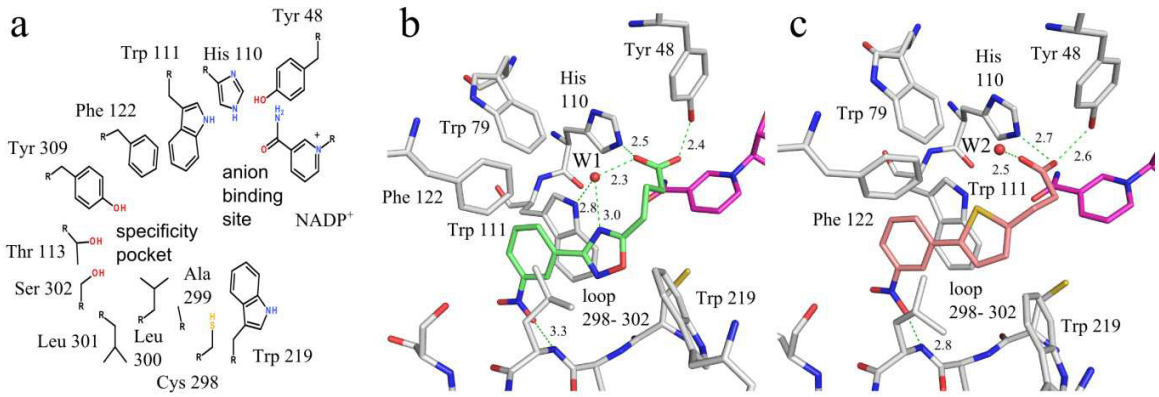



**Figure 4-1** a) Schematic representation of the binding pocket of ALR2. b) Binding mode of **26** (ALR2-**26**, PDB: 2IKG, 1.43 Å).[8] Nitrogen atoms are shown in blue, oxygen atoms in red and sulfur atoms in yellow. Carbon atoms are shown in grey for the protein residues, magenta for $NADP^+$ and in green for **26**. Water molecules which are involved in an H-bond network with the ligand are shown as red spheres. H-bonds are depicted as green dotted lines. This color code is maintained throughout the following representations if not stated otherwise. c) Binding mode of ligand **28** in space group P1 (ALR2-**28**-P1, PDB:3DN5, 1.45 Å).[143] The carbon atoms of the ligand are colored in salmon. The figures 1b) and 1c) were prepared using PyMOL.[36]

It stabilizes the negatively charged group of ALR2 inhibitors, e.g. hydantoin-, succinimide- or carboxylate groups, by H-bonds and a charged interaction to the positively charged cofactor.[38] The specificity pocket shows significant more flexibility. It is bordered to one side by residues which are part of the C-terminal loop. These residues, especially around Leu300, show a high flexibility, which allows the opening of the latter pocket.[37] In closed conformation, the side chain of Leu300 is directed towards Phe122 and blocks the entrance. Due to conformational changes Leu300 turns away which opens the specificity pocket. A closely related protein, human aldehyde reductase (EC 1.1.1.2), does not open a similar subpocket as easily as ALR2 since it is stabilized by a salt bridge in the closed conformation and therefore should not be inhibited.[39] Consequently, ALR2 inhibitors are often decorated with moieties which are able to address this pocket to gain some specificity for ALR2.

A group of carboxylic acid derivatives with two or three aromatic systems was found by a virtual screening campaign in our laboratory which was based on the ultrahigh resolution structure of the ALR2-**25** (IDD594) complex structure (PDB: 1US0).[65] All ligands of this group possess a carboxylate acid head group, a central aromatic moiety ("X" in Figure 4-2) and a terminal aromatic moiety ("Y" in Figure 4-2).

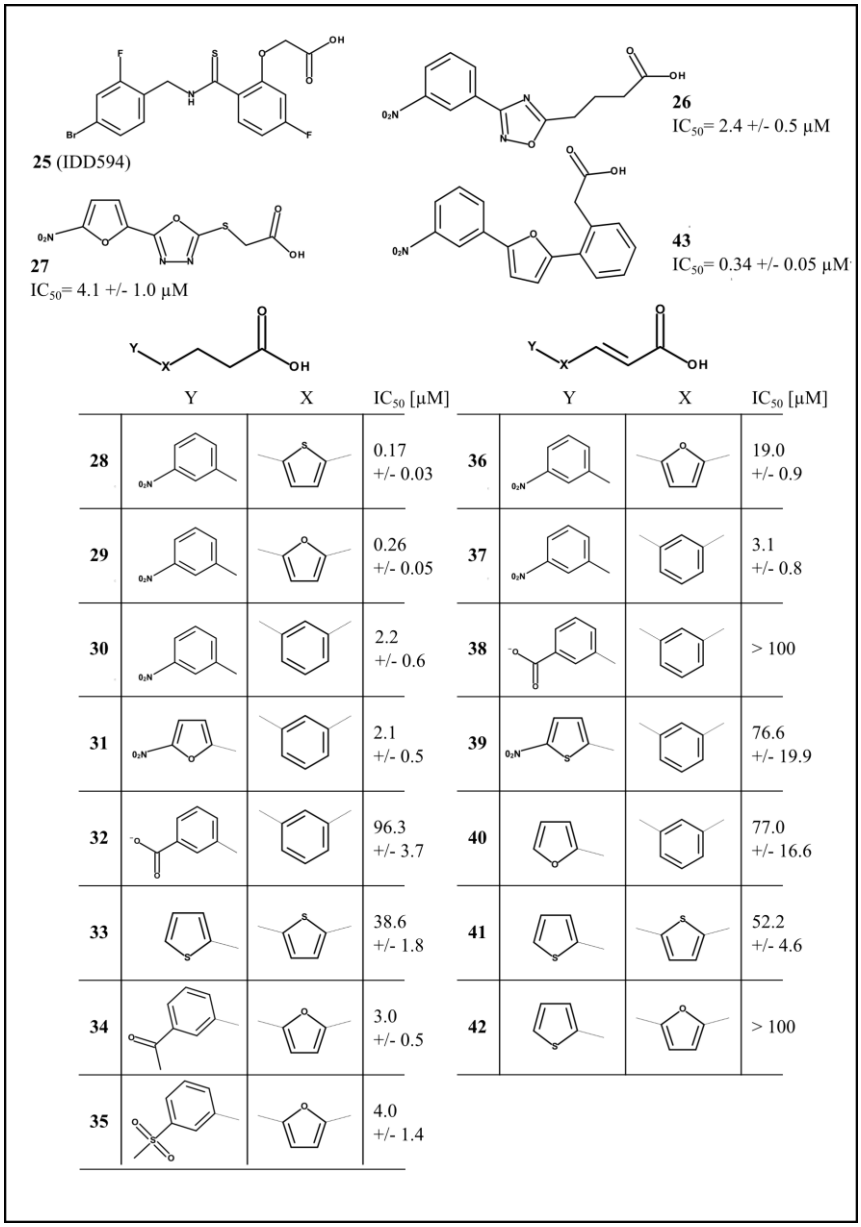



**Figure 4-2** Chemical structures and $IC_{50}$ values of the inhibitors which are investigated in this study. The $IC_{50}$-values for **26**,[65] **27**,[65] **28**,[143] **29**,[143] **34**,[143] **35**,[143] **36**,[143] and **43**[144] were taken from literature.

The carboxylate head group is linked by two (in case of the propionic acid and acrylic acid derivatives) or three carbon atoms (in case of butyric acid derivatives) to the central aromatic moiety. Different substituents are attached to the terminal aromatic moiety. The ligands **26** and **27** were previously characterized crystallographically.[8] Both ligands show $IC_{50}$ values in the low

micromolar range (**26**: 2.4 +/- 0.5 µM, **27**: 4.1 +/- 1.0 µM). In a subsequent optimizing cycle further derivatives were synthesized with the aim of replacing the oxadiazole moiety of **26** and **27** by a thiophene or furan moiety since the heteroatoms of the oxadiazole ring were not involved in polar interactions. This resulted in a series of further compounds between which **28** (0.17 +/- 0.03 µM) and **29** (0.26 +/- 0.05 µM) showed an improved affinity for ALR2. Furthermore the crystal structure of **28** complexed to ALR2 could be determined by cocrystallization in the space group P1.[143] **36** is the side chain analogue of **29** with an acrylic acid head group instead of a propionic acid head group. It showed a significantly decreased affinity (19.0 +/- 0.9 µM) compared to the latter one. Here we present the results of a subsequent optimization cycle. The aromatic moieties and the substituents of the terminal aromatic moiety have been varied to understand their contribution to the affinity and thermodynamic profile. Derivatives containing an acrylic acid function are investigated to further probe the ability of the latter to address the anion binding pocket. Finally, it is evaluated wether the binding pocket of ALR2 can accommodate members of this inhibitor series which comprise a third aromatic moiety. The $IC_{50}$-values are determined. The complex crystal structures of **30**, **31**, **37** and **43** could be determined. Furthermore the complex crystal structure of ALR2-**28** could be determined at high resolution in space group $P2_1$ and is compared with the formerly published structure in space group P1. Additionally the thermodynamic profiles of **26**, **28**, **29**, **30** and **31** could be measured and compared relatively to each other.

## 4.4 Results and Discussion

### 4.4.1 $IC_{50}$-value determination

The $IC_{50}$ values of ligands **26**-**43** are shown in Figure 4-2. **30** ($IC_{50}$: 2.2 +/- 0.6 µM) possesses as **28** and **29** a terminal *meta* nitro phenyl side chain but as a central aromatic moiety it exhibits a phenyl ring instead of a thiophene or a furan moiety. This replacement corresponds to an approximately tenfold decreased affinity compared to the aforementioned which are nearly equipotent. In turn the exchange of the terminal phenyl residue by a furan ring is not accompanied by a significant drop in affinity as shown by the comparison with **31** ($IC_{50}$= 2.1 +/- 0.5 µM). With the exception of **39** ($IC_{50}$= 76.6 +/- 19.9 µM) all ligands possessing a *meta* nitro substituent attached to the terminal aromatic moiety exhibit submicromolar to single digit

micromolar affinity. The favorable contribution to the affinity of the attached *meta* nitro group is in good agreement with previous results.[143] Owing to the obvious importance of this group we decided to probe the influence of an exchange of the nitro group by the isostructural but permanently charged carboxylic acid group. The latter group should be able to form similar interactions in the specificity pocket as the nitro group. Interestingly the $IC_{50}$ value is about fifty times higher for **32** (96.3 +/- 3.7 µM) compared to **30** (2.2 +/- 0.6 µM). Either the carboxylate group forms far less efficient interactions in the specificity pocket or the adopted binding mode of the ligand is entirely different. Since the complex structure of **32** could not be determined due to the low affinity of the ligand a structural validation of the latter hypotheses cannot be provided. The exchange of the propionic acid head group to an acrylic acid function leads to inconsistent results. The propionic acid derivative **33** ($IC_{50}$= 38.6 +/- 1.8 µM) has a slightly higher affinity than its acrylic acid derivative **41** ($IC_{50}$= 52.2 +/- 4.6 µM) which is in agreement with former studies.[143] On the other hand **37** ($IC_{50}$= 3.1 +/- 0.8 µM) which possesses an acrylic acid head group shows nearly the same affinity as **30** ($IC_{50}$= 2.2 +/- 0.6 µM) which comprises a propionic acid head group instead. The explanation for these deviating results appears quite controversal. As will be explained later a change of the configuration of the double bond from *E*- to *Z*- for **37** is found while analyzing the corresponding complex crystal structure with ALR2. It could not be resolved whether this configurational transition occurs for all the ligands containing an acrylic acid head group prior to binding to ALR2. In case the configurational transition takes place to deviating extent for the different ligands investigated here, this might explain why in some cases the replacement of the propylene side chain by the double bond comes with a decreased affinity and in other cases not. A similar drop of affinity due to the exchange of the *meta* nitro group by a *meta* carboxylate group as described above for the ligands bearing a propionic acid head group is observed for a member with an acrylic acid head group: **37** with a terminal nitro group ($IC_{50}$= 3.1 +/- 0.8 µM) exhibits much higher affinity compared to the carboxylate derivative **38** ($IC_{50}$ >100 µM). **43** contains a third aromatic ring system. The additional phenyl portion and the fixed headgroup arrangement are well tolerated by ALR2 since the $IC_{50}$ value rates the compound as high affinity binder (0.34 +/- 0.05 µM). This finding suggests that the ligand adopts a binding mode with ALR2 similar to the other highly potent members of this ligand series despite its spatially more demanding shape. This hypothesis will be further investigated by crystallography.

### 4.4.2 Crystallographic determination of the binding modes of 30, 31, 37 and 43

#### 4.4.2.1 Similarities in the binding modes of 26, 28, 30, 31, 37 and 43

The overall binding modes of **26**, **28**, **30**, **31**, **37** and **43** are similar. Figure 4-3a shows the superposition of the inhibitor molecules obtained by mutual alignment of the complex crystal structures. Two anchoring points can be defined which correspond to the two important interactions accomplished by the potent ligands in the binding site. On the one hand it can be observed that the carboxylic acid function aligns almost perfectly for **30**, **31**, **37** and **43** while it is orientated slightly different for **26** and **28**. This part of the ligands binds to the anion binding site. On the other hand the position of the terminal nitro group matches exactly for all studied ligands. The remaining ligand scaffolds arrange in a way to accomplish these key interactions. The alignment of all complex structures shows that all ligands adopt a very similar spatial occupancy (Figure 4-3a, Figure 4-4a). All inhibitors open the specificity pocket with their terminal nitro phenyl moiety. The $IC_{50}$ values show that the introduction of the nitro group leads to a significantly higher binding affinity. As shown in Figure 4-3b and Figure 4-4c in ALR2-**30** the nitro group forms via one of its oxygen atoms an H-bond to the backbone nitrogen of Leu300 (3.0 Å). Similar H-bonds are established by **26** (3.3 Å, Figure 4-1b), **28** (3.0 Å in space group $P2_1$, Figure 4-4b), **31** (3.1 Å, Figure 4-4d), **37** (3.1 Å, Figure 4-4e) and **43** (3.1 Å, Figure 4-4f). Interestingly in syn direction the distance between the two oxygen atoms of the nitro group of the ligands and the carbon atoms of the phenyl ring of Tyr309 are surprisingly short in all mentioned ALR2-ligand complexes. As shown exemplarily in Figure 4-3b for ALR2-**30** the distances between the oxygen atoms of the nitro group of **30** to the carbon atoms of Tyr309 are only 3.3 Å and 3.4 Å, respectively. It was discussed previously that these short distances indicate so-called non-classical hydrogen bonds between the oxygen atoms of the nitro group and the hydrogen atoms attached to the phenyl ring.[8] No water molecules are directly bound to **30**, **31**, **37** and **43**. The complex structures of **30** and **37** were obtained at pH 5.5, the complex structures of **28**, **31** and **43** at pH 8.0. At both pH values the carboxylic acid function is presumably deprotonated and should carry a negative charge. As known from a former study with the carboxylate-type ligand **25** (IDD594, PDB: 1US0)[40] the cofactor is bound in its oxidized form and is thus present with a positive charge. When a similar situation is assumed in the here described ligand series a charge-

assisted interaction between the presumably negative charged carboxylate group of the ligands and the positively charged $NADP^+$ will be established.

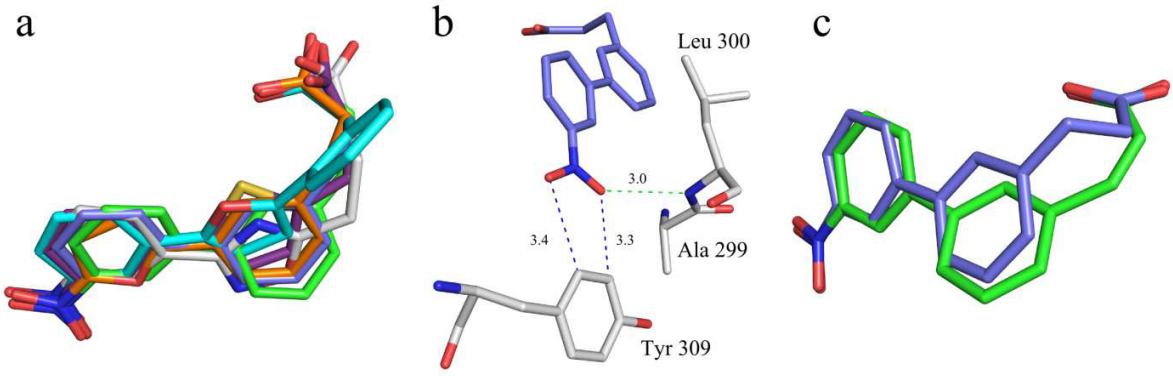



**Figure 4-3** a) Alignment of the ligands only. Carbon atoms for **26** are shown in grey, for **28** (in space group $P2_1$) in purple, for **30** in blue, for **31** in orange, for **37** in light green and for **43** in cyan. The nitro groups of all ligands align very well. The carboxylate groups of **26** and **28** adopt a kinked position compared to the carboxylate groups of **30**, **31**, **37** and **43**. b) Interactions of the nitro group of **30** with Tyr309 and Leu300. Non-classical H-bonds are shown as blue-dotted lines. Non-classical hydrogen bonds are formed between the oxygen atoms of the nitro group of **30** and the carbon atoms of the phenyl ring of Tyr309. Furthermore, an H-bond between one of the oxygen atoms of the nitro group to the amide bond of Leu300 is established. The other inhibitors are forming similar interaction since their nitro group is positioned likewise. c) Comparison of the binding mode of **30** and **37**. The carbon atoms of **30** are colored in blue and of **37** in green. The central phenyl ring of **37** is spatially shifted compared to the corresponding ring in **30**. Presumably, the reduced flexibility of **37** compared to **30** forces **37** into this orientation. Figures were prepared using PyMOL.[36]

### 4.4.2.2 Binding modes of 26 and 28

The crystal structures of **26** (PDB: 2IKG, 1.43 Å) and **28** (PDB: 3DN5, 1.45 Å) complexed to ALR2 were described previously[8, 143] (Figure 4-1b-c). The carboxylic acid functions of both ligands are bound to the anion binding pocket. In ALR2-**26** it forms H-bonds to Tyr48OH (2.4 Å) with one of its oxygen atoms and to His110NE2 (2.5 Å) with the other oxygen atom. Among the investigated compounds studied in the present series, **26** establishes the shortest H-bond to Tyr48 (2.4 Å). Furthermore, the latter oxygen of the carboxylic group is not involved in any further interactions suggesting this contact as a very strong H-bond. **26** possesses a butyric acid head group.

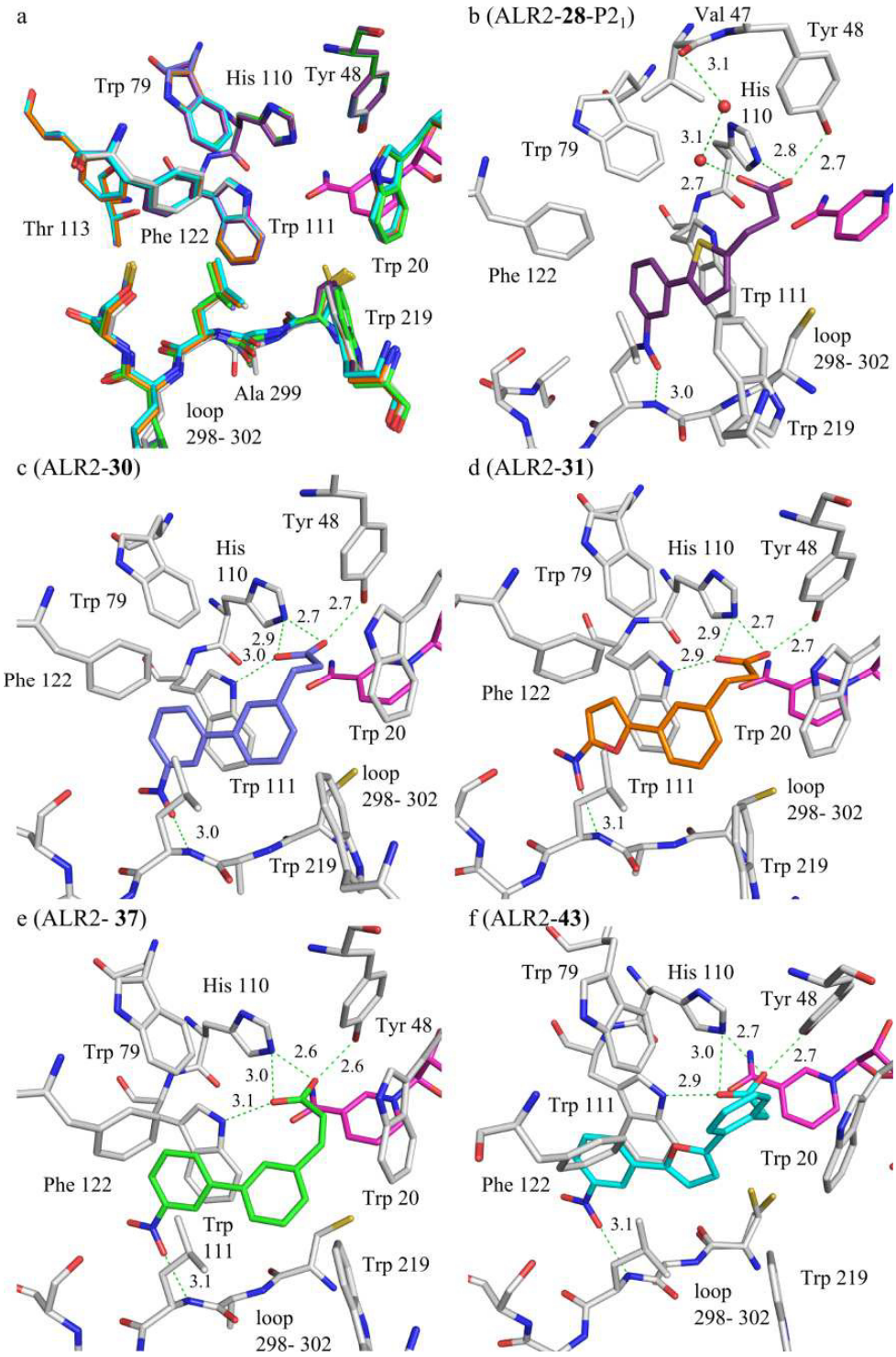



**Figure 4-4** Detailed figure caption on next page.

**Figure 4-4 Stick representation of the binding pocket of ALR2 complexed with 28, 30, 31, 37 and 43**. a) Superposition of the binding pockets of the complexes ALR2-**26**, ALR2-**28**-$P2_1$, ALR2-**30**, ALR2-**31**, ALR2-**37** and ALR2-**43**. The carbon atoms of the protein residues of ALR2-**26** are colored in grey, the carbon atoms of the residues of the remaining protein-ligand complexes are colored corresponding to the colors of the carbon atoms of the ligands in b)-f). All residues of the binding site align very well. In consequence, all the ligands induce a very similar binding site conformation. b) Binding mode of **28** in space group $P2_1$ (ALR2-**28**-$P2_1$, PDB: 4YU1, 1.02 Å). The carbon atoms of the protein residues are shown in grey and of $NADP^+$ in purple. This will be maintained during the following representations. The carbon atoms of **28** are shown in purple. c) Binding mode of **30** (PDB: 4PR4, 1.06 Å). Carbon atoms of **30** are colored in blue. d) Binding mode of **31** (PDB: 4PRR, 1.01 Å). Carbon atoms of **31** are colored in orange. e) Binding mode of **37** (PDB: 4NKC, 1.12 Å). Carbon atoms of **37** are colored in green. f) Binding mode of **43** (PDB: 4PRT, 0.96 Å). Carbon atoms of **43** are colored in cyan.

This long linker between the carboxylic acid function and the central aromatic moiety prevents a direct H-bond between the carboxylic acid group and Trp111 which is found for many of the other ALR2 inhibitors (Figure 4-4).[40, 127] Instead a water molecule (W1 in Figure 4-1b) is captured which mediates the contact to Trp111NE1 (2.8 Å) and to one of the oxygen atoms of the carboxylate group of **26** (2.3 Å). The water molecule is further stabilized by an H-bond to one of the nitrogen atoms of the central oxadiazol ring of the ligand (3.0 Å).[8] The crystal structure was obtained at pH 5.0 at which the carboxylic acid group is assumed to be mainly deprotonated. A charge-assisted interaction between the negatively charged carboxylic acid group and the positively charged $NADP^+$ is detected (2.8 Å). The ALR2-**26** structure was obtained by soaking preformed ALR2 crystals, crystallized in space group $P2_1$, in a solution containing **26**. In contrast the structure of **28** complexed to ALR2 (in the following referred to as ALR2-**28**-P1) was obtained by co-crystallization. In the latter case the ternary complex of ALR2, $NADP^+$ and **28** crystallized in space group P1.[143] One of the oxygen atoms of the carboxylate group of **28** forms H-bonds to Tyr48OH (2.6 Å) and His110NE2 (2.7 Å). Furthermore, the carboxylic acid group forms with the latter oxygen atom the charge-assisted interaction to $NADP^+$. Similarly as in the ALR2-**26** structure the carboxylate group of **28** does not form a direct H-bond to Trp111. Instead one of the oxygen atoms of the carboxylate group of **28** establishes an H-bond to a water molecule (W2 in Figure 4-1c). To better understand the role of this water molecule and to estimate whether its position is influenced by the space group change observed for the latter crystal structure an additional structure determination was attempted. This time an uncomplexed ALR2 crystal, obtained in space group $P2_1$, was soaked in a saturated solution of **28** in TRIS buffer at pH 8.0 (in the following referred to as ALR2-**28**-$P2_1$).

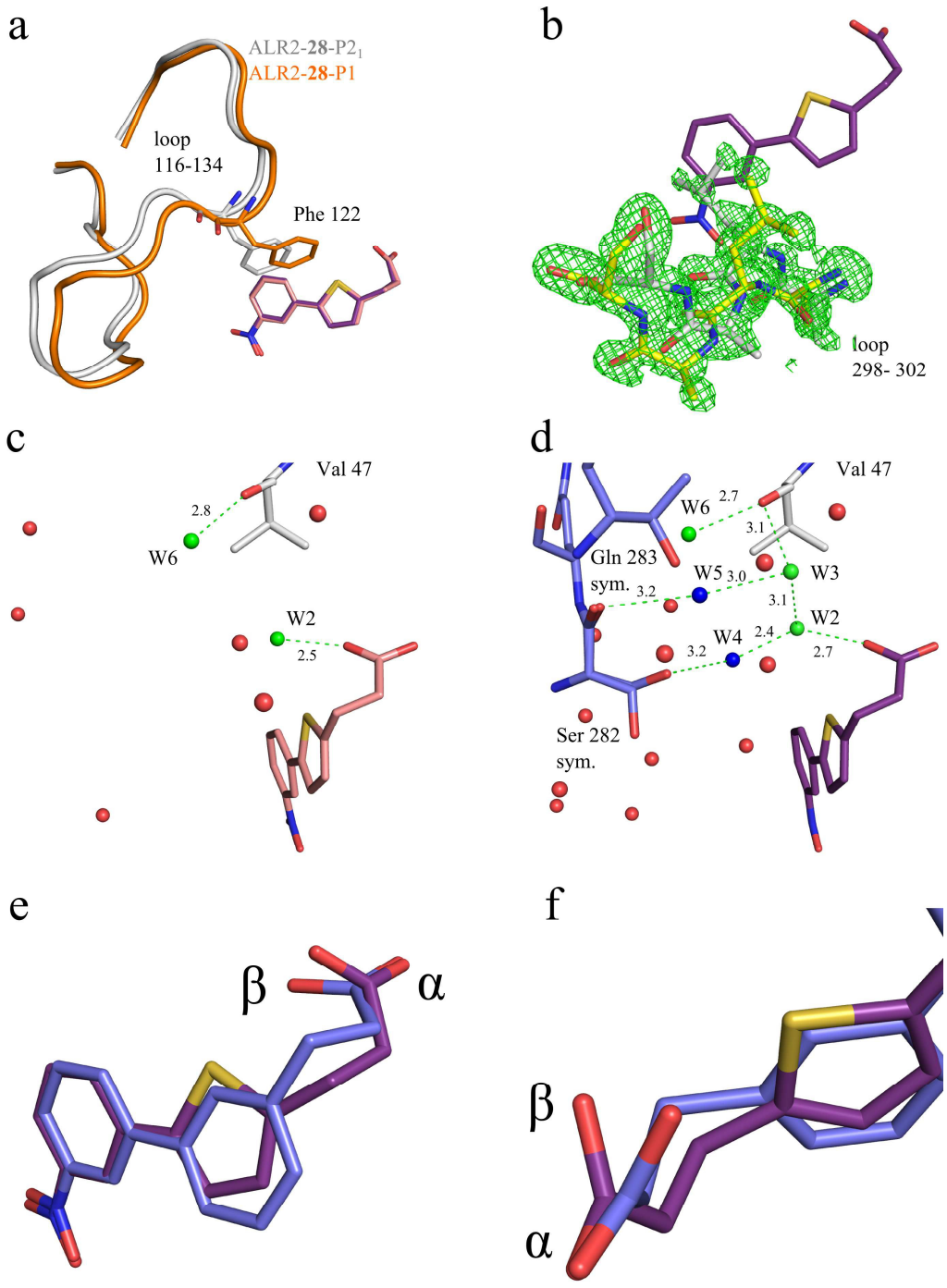



**Figure 4-5** Detailed figure caption on next page.

**Figure 4-5** Comparison of the complex crystal structures of **28** in space groups P1 (ALR2-**28**-P1) and $P2_1$ (ALR2-**28**-$P2_1$). a) Alignment of ALR2-**28**-P1[143] and ALR2-**28**-$P2_1$. The carbon atoms of the residues 116-134 are shown in grey for ALR2-**28**-$P2_1$ and in orange for ALR2-**28**-P1. The carbon atoms of **28** are colored in purple in ALR2-**28**-$P2_1$ and in salmon in ALR2-**28**-P1. In ALR2-**28**-$P2_1$ the loop consisting of residues 116-134 is shifted compared to ALR2-**28**-P1, which is followed by a different orientation of the side chain of Phe122. The spatial orientation of **28** aligns perfectly well in both structures and thus the adopted binding pose is not influenced by the space group. b) Representation of the $F_o$-$F_c$ difference density map at 3.0 σ after five cycles of simulated annealing for residues 298-302 in ALR2-**28**-$P2_1$. The carbon atoms of the ligand are shown in purple, the carbon atoms of the protein residues corresponding to the specificity pocket in open conformation, are depicted in yellow and the ones corresponding to the closed conformation are displayed in grey. The occupancy of the ligand was refined to 70%. The specificity pocket is presumably only opened when the ligand has bound. Accordingly 30% of the the protein residues 298-302 are found in the conformation which corresponds to the closed conformation of the specificity pocket. c) Water structure in ALR2-**28**-P1.[143] The water molecules W2 and W6 which are in H-bonding distances to **28** and Val 47 are shown as green spheres, further water molecules are shown as red spheres. W2 does not form further interactions apart from the H-bond to **28**. d) Water structures in ALR2-**28**-$P2_1$. The carbon atoms of the protein residues are colored in grey, those which belong to the protein residues of the next symmetry mate are shown in blue. The water molecules W2, W3 and W6 which are in H-bond distances to **28** and Val47 are shown as green spheres. The water molecules W4 and W5 which are stabilized by the amino acid residues of the next symmetry mate are shown as blue spheres. Further water molecules are shown as red spheres. e) and f): Alignment of ALR2-**30** and ALR2-**28**-$P2_1$. A superimposition of the two ligands is shown. Due to the different central aromatic ring systems the positions of the carboxylate groups of both ligands vary. The α- oxygen atoms align well but the positions of the β- oxygen atoms differ.

The binding modes of **28** in both structures align perfectly well (Figure 4-5a). Koch et al. described previously the influence of crystal packing on the positioning of the loop consisting of the amino acids 116-134. They found that Lys119 and Phe121 form H-bonds to Gln283 of the next symmetry mate in the $P2_1$ crystal packing. In P1 these residues are not involved in any crystal contacts.[145] This leads to a shift of the loop consisting of residues 116-134 in crystals which were obtained in space group $P2_1$ with respect to those obtained in space group P1. This affects mainly the position of the side chain of Phe122. In the present comparison of ALR2-**28**-P1 and ALR2-**28**-$P2_1$ we observe the same finding (Figure 4-5a). Surprisingly, in ALR2-**28**-$P2_1$ the ligand showed only an occupancy of 70%. After initial placement of the ligand into the electron density the structure was refined. Afterwards still some $F_o$-$F_c$ difference electron density at the 3.0 σ level in agreement with the closed conformation of the specificity pocket is visible. The coordinates of the protein residues 298-303 which are mainly involved in the flexibility and opening of the specificity pocket were built in two conformations into the density. The occupancy of conformation of the latter residues corresponding to the open conformation of the specificity pocket was fixed to 70% in agreement with the occupany of **28** which induces through binding the opening of the pocket. The closed conformation of the specificity pocket was fitted into the difference electron density at the 3.0 σ level. The refinement resulted in a $2F_o$-$F_c$ density

electron map at 1.0 σ which was in good agreement with the coordinates of the closed conformation. To exclude a major model bias, both conformations of the residues 298-303 were deleted from the model and five cycles of simulated annealing were performed. The resulting $F_o$-$F_c$ difference electron density map at 3.0 σ is shown in Figure 4-5b. It shows density both for the specificity pocket in closed and in open state. Consequently, ALR2-**28**-$P2_1$ was deposited with the specificity pocket modeled in both conformations. The position of water W2 which was already visible in ALR2-**28**-P1 (Figure 4-5c) could be confirmed in the $P2_1$ complex. Interestingly here it is connected to a further water molecule (W3 in Figure 4-5d, 3.1 Å). Furthermore both W2 and W3 are stabilized by H-bonds to the water molecules W4 (2.4 Å) and W5 (3.0 Å), respectively. Both W4 and W5 are stabilized themselves by H-bonds to Ser282 of the next symmetry mate. W4 forms an H-bond to the OH group of Ser 282 (3.2 Å) and W5 to the backbone oxygen (3.2 Å) of the same residue. Whether the occupancy of W3 in ALR2-**28**-$P2_1$ is mainly stabilized by the remote contacts to the next symmetry mate remains difficult to answer. In both structures another water molecule (W6) is in H-bond contact to the carbonyl oxygen atom of Val47 (2.8 Å and 2.7 Å, respectively). In ALR2-**28**-P1 the location of W3 remains unoccupied but no sterical repulsion with adjacent protein residues or other water molecules would argue against the presence of a water molecule. The higher resolution of the ALR2-**28**-$P2_1$ (1.02 Å compared to 1.45 Å) structure might be a reason why W3 was actually not detected in the P1 structure. As W3 is in additionally connected to the backbone oxygen atom of Val47 (3.1 Å ) in ALR2-**28**-$P2_1$ the latter structure shows that **28** uses W2 and W3 to establish a water network to form a bridge to Val47.

#### 4.4.2.3 Binding mode of the carboxylic acid head group in the anion binding pocket of ligands 30, 31, 37 and 43.

The carboxylic acid head groups of **30**, **31**, **37** and **43** bind as assumed in the anion binding pocket. They all form H-bonds between one of the oxygen atoms of the carboxylate group and Tyr48OH (2.6-2.7 Å) and His110NE2 (2.6-2.7 Å) while the other oxygen atom of the carboxylate group forms H-bonds to His110NE2 (2.6-3.0 Å) and Trp111NE1 (2.9-3.1 Å). Thus, in contrast to ALR2-**26** and ALR2-**28**-$P2_1$, a direct contact to Trp111 is formed. All ligands are involved in the already discussed charge-assisted interaction between one of the oxygen atoms of the deprotonated carboxylic acid group and the positively charged $NADP^+$ (2.9-3.0 Å).

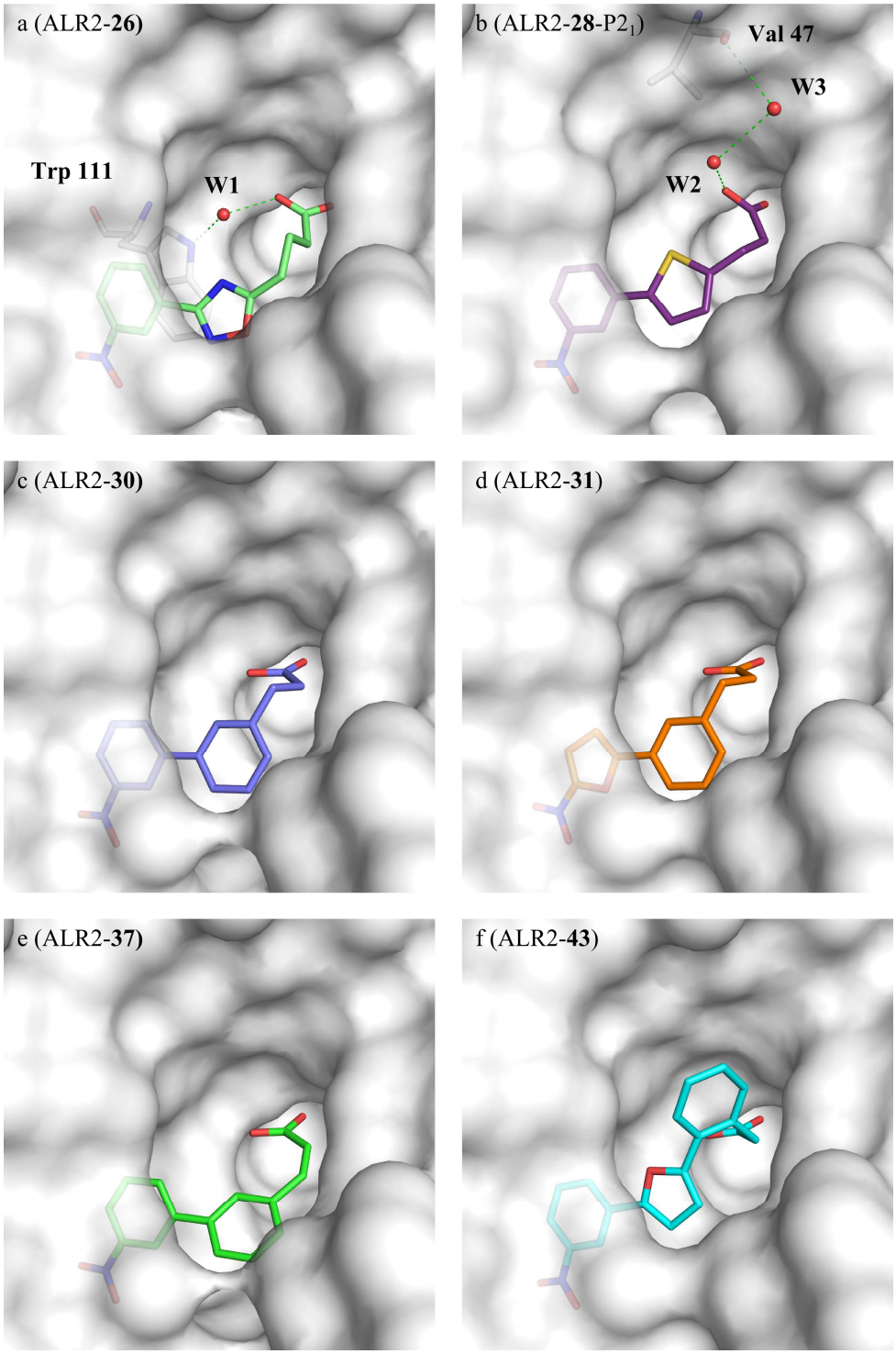



**Figure 4-6** Detailed figure caption on next page.

**Figure 4-6 Surface representation of the binding pocket of ALR2 complexed with ligands 26, 28, 30, 31, 37 and 43.** The solvent-accessible surface of the protein is shown in grey. Only water molecules are shown which are involved in an H-bond network with the ligand. All ligands address with their *meta* nitro phenyl moiety the specificity pocket. All structures are shown from the same viewing direction. a) Binding mode of **26** (ALR2-**26**, PDB: 2IKG, 1.43 Å).[8] The carbon atoms of ligand **26** are colored in green. The water molecule W1 mediates an H-bond between one of the oxygen atoms of **26** and the Trp111NE1. b) Binding mode of **28** in space group $P2_1$ (ALR2-**28**-$P2_1$, PDB: 4YU1, 1.02 Å). The carbon atoms of **28** are shown in purple. Two water molecules (W2 and W3) mediate an H-bond network between one of the oxygen atoms of **28** and the OH group of Ser282. c) Binding mode of **30** (PDB: 4PR4, 1.06 Å). Carbon atoms of **30** are colored in blue. d) Binding mode of **31** (PDB: 4PRR, 1.01 Å). Carbon atoms of **31** are colored in orange. e) Binding mode of **37** (PDB: 4NKC, 1.12 Å). Carbon atoms of **37** are colored in green. f) Binding mode of **43** (PDB: 4PRT, 0.96 Å). Carbon atoms of **43** are colored in cyan.

#### 4.4.2.4 Binding mode of 30

The structure of **30** complexed to ALR2 was refined to a resolution of 1.06 Å. **30** possesses similar to **37** and **43** a terminal *meta* phenyl moiety which addresses the specificity pocket. The central aromatic moiety consists here of a phenyl ring which is connected via a linker of two carbon atoms to the carboxylic acid function. The terminal phenyl ring forms a face-to-face π-π-stacking to Trp111 (3.6 Å) in the specificity pocket. The central phenyl ring and the C2-linker establish hydrophobic interactions to Trp20 (3.5 Å), Trp219 (3.6 Å), Cys298 (3.8 Å), Leu300 (3.5 Å) and Trp111 (3.7 Å, Figure 4-4 c).

#### 4.4.2.5 Binding mode of 31

The complex crystal structure ALR2-**31** was refined to a resolution of 1.01 Å. **31** possesses in comparison to **30** a terminal 5-nitro-2-furan ring. Similar to **30** it contains a central phenyl ring and a C2-linker to the carboxylic acid function. The binding modes of **30** and **31** are very similar. The same hydrophobic contacts of the central part of **31** consisting of the phenyl ring and the C2 linker are established as for **30**. In detail, these are hydrophobic interactions to Trp20 (3.6 Å), Trp219 (3.7 Å), Cys298 (3.8 Å), Leu300 (3.5 Å) and Trp111 (3.7 Å). The oxygen of the furan ring of **31** does not form any H-bonds. The backbone of Leu300 forms an H-bond to the nitro group of **31** (3.1 Å) and is too far away (3.9 Å) to form another H-bond with the furan ring of **31**. The furan ring makes a face-to-face π-π-stacking to Trp111 (3.8 Å) in the specificity pocket. Regarding the required prize to spend for desolvation effect it appears not favorable to use a furyl moiety to address the specificity pocket. The oxygen atom of the furan ring has to be desolvated when bound to the protein but as the structure shows it is not compensated by forming any polar interactions with the protein (Figure 4-4d).

#### 4.4.2.6 Binding mode of 37

The complex crystal structure ALR2-**37** was refined to a resolution of 1.12 Å. **37** is the only compound of this series with an acrylic acid head group for which a complex crystal structure could be determined. As a terminal group it posseses as **30** and **43** a *meta* nitro phenol moiety. **37** was initially synthesized as the *E*-isomer. Suprisingly, in ALR2-**37** only the *Z*-isomer is visible. It was shown by $^{1}$H-NMR analysis that the geometry of the double bond of **37** is liable to isomerization (see Section 4.6.2). This surprising observation clearly indicates that ALR2 strongly favors the ligand with the double bond in *Z*-configuration at this position. The rigidification of a ligand with a double bond in the required binding site geometry will enhance the free energy of binding. The entropic portion to binding will be favorably influenced since a more rigidified ligand loses less degrees of freedom compared to a more flexible ligand. This experience can be used as a strategy for the development of further ALR2-inhibitors. We considered the possibility to synthesize **37** with the double bond in *Z*-configuration. However, since the configuration of this bond seems not to be stable in solution we discarded this attempt. The terminal phenyl ring makes in the specificity pocket a face-to-face π-π-stacking to Trp111 (3.7 Å). The central part of **37** which consists of the phenyl ring and the *Z*-configurated double bond C2 linker makes hydrophobic contacts to Trp20 (3.5 Å), Trp219 (3.4 Å), Leu300 (3.5 Å), Cys298 (3.6 Å) and Trp111 (3.5 Å). Since the molecule possesses a limited flexibility due to its double bond, the central phenyl ring deviates from the plane which is defined by the corresponding phenyl rings of **30** and **31**. The superposition of **30** and **37** is shown in Figure 4-3c. It is obvious that **37** binds to ALR2 in a way that the interactions of the carboxylic acid group and of the *meta* nitro group are possible adopting a geometry highly similar to that of **30**. In contrast the central part of **37** shifts compared to **30** (Figure 4-4e).

#### 4.4.2.7 Binding mode of 43

The crystal structure of **43** complexed to ALR2 was refined to a resolution of 0.96 Å. **43** is the only member of this group of aldose reductase inhibitors which exhibits a third aromatic moiety and for which crystal structure determination was successful. An additional phenyl ring serves as a linker between the central furan ring and the acetic acid function which binds in the anion binding site. **43** possesses as **30** and **37** a terminal *meta* nitro phenyl moiety. The oxygen atom of the central furan ring of **43** is not involved in any H-bonds and it is directed towards the

hydrophobic residues Trp79 and Phe122. This oxygen atom cannot compensate for shedding its solvation shell by any newly formed polar contacts to the protein. The phenyl ring next to the acetic acid group forms hydrophobic contacts to Trp79 (4.1 Å), Tyr48 (3.7 Å) and Trp20 (4.2 Å). A π-π- stacking is here not observed, nevertheless the phenyl moiety should be well accommodated in this predominantly aromatic environment. The furan ring forms hydrophobic interactions to Leu300 (3.4 Å) and Trp111 (3.3 Å). In the specificity pocket the terminal phenyl ring undergoes a face-to-face π-π-stacking to Trp111 (3.5 Å, Figure 4-4f).

### 4.4.3 Further crystallographic experiments

The initial aim of this study was to elucidate the complex crystal structures of numerous members of this ligand series and measure the corresponding thermodynamic signature. At the beginning of this project ALR2 complex structures were either obtained by cocrystallization or by ligand soaking in citrate buffer. The complex crystal structures of ALR2-**26**, ALR2-**27** and ALR2-**28**-P1 were obtained under these conditions.[8, 143] One example was known for which it was possible to cocrystallize the protein with an inhibitor in TRIS buffer at pH 8.0[121] but these conditions were difficult to reproduce. Unfortunately, in none of the various collected datasets of cocrystallized or soaked ALR2 crystals in citrate buffer sufficient difference electron density for the accomodation of the bound ligand molecule could be found. The description of a new crystallization condition in MES-buffer, pH 5.5, by Cousido-Siah et al.[101] was a very valuable asset for our crystallographic study. Cousido- Siah et al. discussed the absence of citrate in the MES condition as crucial since citrate is known to be an aldose reductase inhibitor itself[38] and thus it always competes with a ligand to be bound in the crystallographic analysis. Plenty of datasets of ALR2 crystals either cocrystallized or soaked under the citrate condition were collected. In fact often not enough difference electron density for a bound ligand was visible whereas still some difference electron density indicated the presence of a carboxylic acid group in the anion binding pocket. This could result from the occupancy with disordered molecules of citrate as described in literature[34] and supports the hypothesis that citrate is blocking the binding of the inhibitors. Since it required some time to establish the crystallization in MES in our laboratory an alternative approach was tested. The protein was crystallized under the known condition in citrate buffer at pH 5.0. Afterwards the crystals were transferred into several droplets

in a step-wise fashion which contained the initial citrate condition with 25% (w/v) PEG 6000 at pH 5.0 in decreasing concentration and the MES condition with 25% (w/v) PEG 6000 at pH 5.5 in increasing concentrations. The latter solution was saturated with the inhibitor in order to increase its concentration. For several ligands this protocol was applied. In fact, this procedure gave access to the complex structures of **30** and **37**. Still in several crystals soaked with other ligands of this series, for example **31**, not sufficient difference electron density for the bound ligand was detected. In the mean time the crystallization of ALR2 under the MES conditions was established in our laboratory. To our surprise a dataset of the *holo*-enzyme of this condition still showed density for a carboxylic acid function in the anion binding site. Since the protein had never been in contact with citrate there must some other molecule be present showing a similar functional group topology. Presumably the difference electron density is created by another unknown carboxylic acid, obviously picked-up by the protein during the expression or purification. However, since the exchange of citrate to MES as buffer compound seems not to be the essential discriminating factor, the modified pH value of the applied conditions attracted our attention. While at pH 5.0 or 5.5 the carboxylic acid function of the studied ligand molecules could still be partly protonated, it should be fully deprotonated at more basic pH values. Thus the soaking experiments described above were repeated. This time the droplets consisted of decreasing concentrations of citrate buffer with 25% (w/v) PEG 6000 at pH 5.0 and increasing concentrations of TRIS buffer with 25% (w/v) PEG 6000 at pH 8.0, saturated with the inhibitor. In fact **31** did successfully bind under this condition to ALR2. Unfortunately it was the only ligand for which these new soaking conditions brought the expected achievement. Indeed also the complex structure of **43** could be successfully determined in TRIS buffer. However it is possible that this ligand would also bind under different conditions since it possesses a high potency. Interestingly, using the latter described soaking procedure **29**, **34** and **35** lead to datasets where in addition to the missing $F_o$-$F_c$ difference electron density for the ligand molecule also the protein residues next to Leu300 were disordered and thus not visible in the density. Apparently these ligands are not binding in an ordered fashion under these conditions and are so far only able to disturb the crystal packing in this area of the protein. Two factors should be mentioned that may explain the pH dependency of the binding of **30**, **31** and **37** to ALR2 crystals. At higher pH a larger portion of ligand molecules will exhibit a deprotonated carboxylic acid function. This

should increase the binding since the negative charge enhances the binding to the anion binding pocket which contains the positively charged $NADP^+$. On the other hand the deprotonated carboxylate group should lead to higher solubility of the compounds. Since saturated solutions of the inhibitors are used under the soaking buffer conditions the basic pH value will lead to higher concentrations of the ligands which facilitate the binding to ALR2. To further increase the solubility of the compounds up to 10% (v/v) DMSO was added to the soaking solution, but this did not further increase the hit rate. However, these experiments indicate that the success of protein soaking experiments can be improved by careful choice of the soaking conditions.

### 4.4.4 Thermodynamic characterization of 26 and 28-31.

All ITC measurements were performed in HEPES buffer, pH 8.0 (Figure 10-2 in the Appendix). Steuber et al. could show that the binding of carboxylate-type ALR2 inhibitors, among them **26**, is accompanied by an entrapment of 0.8-0.9 protons due to the protonation of Tyr48.[42] The thermodynamic profile of **26** was measured with a MCS-ITC instrument of MicroCal. The reported data for the titration in HEPES buffer were $\Delta G^0$= -35.4 kJ/mol, $\Delta H^0$= -25.6 kJ/mol and $-T\Delta S^0$=-9.8 kJ/mol.[8] The values corrected for the protonation step were $\Delta G^0$= -35.4 kJ/mol, $\Delta H^0$= -45.1 kJ/mol and $-T\Delta S^0$= +9.7 kJ/mol[42] ($-T\Delta S^0$ was calculated from the reported data using the equation $-T\Delta S^0 = \Delta G^0 - \Delta H^0$). Due to the small structural variations we assume a similar protonation behavior for all ligands of which the thermodynamic profile will be discussed in the following. Since the aim of this study was to compare the different ligands relative to one another, we relinquished to perform the measurement in different buffer systems to correct the enthalpy for the constantly superimposed protonation step. The thermodynamic data are shown in Figure 4-7 and in Table 4-1. Representative thermograms and the fitted regression curves are shown in Figure 10-2 in the Appendix. Since all new compounds were measured with an $ITC_{200}$ we repeated the measurements for **26** with the latter instrument. The enthalpy could be reproduced whereas the $K_D$-value and thus the free energy of binding and the entropic term data deviate slightly more ($\Delta G^0$= -31.5 kJ/mol, $\Delta H^0$= -27.2 kJ/mol and $-T\Delta S^0$= -4.3 kJ/mol[8]). For **37** only ITC curves of poor quality were obtained. As explained above the ligand changed its configuration from *E*- to *Z*-conformer prior to binding to ALR2. This process occurs apparently too slowly to be measured by an ITC experiment. Since the change of configuration could be

induced by UV light we exposed a solution of **37** overnight by UV light. However, the following ITC measurement resulted again in a curve of insufficient quality. Presumably still a significant portion of **37** was present in *E*-configuration which disturbed the measurement.

**Table 4-1 $IC_{50}$- values and thermodynamic data for compounds 26, 28, 29, 30 and 31**

| | $IC_{50}$ [µM] | $K_D$ [µM] | $\Delta G^0$ [kJ/mol] | $\Delta H^0$ [kJ/mol] | $-T*\Delta S^0$ [kJ/mol] |
|---|---|---|---|---|---|
| **26** | 2.4 ± 0.5[8] | 3.1 ± 0.3 | -31.5 ± 0.3 | -27.2 ± 0.7 | -4.3 ± 0.9 |
| **28** | 0.17 ± 0.03[143] | 0.82 ± 0.1 | -34.8 ± 0.1 | -33.5 ± 0.1 | -1.3 ± 0.0 |
| **29** | 0.26 ± 0.05[143] | 2.2 ± 0.1 | -32.3 ± 0.1 | -23.4 ± 0.1 | -8.9 ± 0.1 |
| **30** | 2.2 ± 0.6 | 1.7 ± 0.1 | -32.9 ± 0.0 | -27.4 ± 0.1 | -5.5 ± 0.1 |
| **31** | 2.1 ± 0.5 | 1.9 ± 0.1 | -32.6 ± 0.1 | -24.4 ± 0.4 | -8.2 ± 0.4 |

The thermodynamic signature of the binding of a ligand to a protein is composed of different factors. Each polar interaction e. g. H-bond or charge-assisted contact contributes favorably to the enthalpic portion. It can be assumed that especially the charge-assisted interaction between the negatively charged carboxylate group of this ligand series and the positively charged $NADP^+$ has a strong favorable influence on the enthalpic portion. The binding pocket of ALR2 contains several hydrophobic residues (e.g. Trp20, Trp111, Phe122, Pro218, Ala299, Leu300) even with the specificity pocket in its closed state. Prior to the accommodation of a ligand, water molecules will likely fill the binding pocket. These water molecules can be assumed to be entropically unfavored compared to the situation in the bulk water phase and their release should lead to a favorable signal in the entropic portion of the thermodynamic profile. This would be in line with the classical hydrophobic effect.[146-148] This general consideration should be valid for all here investigated ligands since they all undergo van der Waals interactions with the non-polar residues of the binding pocket and consequently are displacing water molecules from this hydrophobic environment. Since no highly resolved structure of the *holo*-enzyme of ALR2 is available it is impossible to estimate how many water molecules are displaced by the ligand. In terms of their molecular volume the studied ligands fall into the same range (221.0-239.4 $A^3$, $\Delta V$= 18.4 $A^3$, calculated with Molinspiration property engine)[149] making the hypothesis of the replacement of similar amount of water molecules at least reasonable.

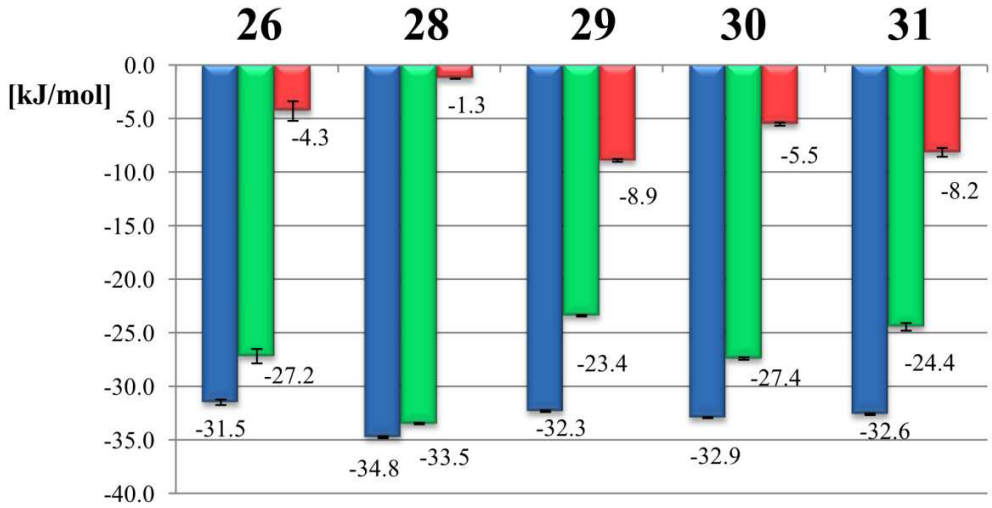



**Figure 4-7 Thermodynamic profiles of 26, 28, 30, 31 and 29 measured in HEPES.** The free energy of binding (ΔG°) is shown as blue bars, the enthalpy (ΔH°) as green bars and the entropic contribution (-T*ΔS) as red bars.

In the following, the differences in the thermodynamic profiles of the ligands will be discussed with respect to their corresponding binding modes. In HEPES buffer the thermodynamic profile of **30** is dominated by the enthalpic part (-27.4 kJ/mol) and has only a small favorable entropic contribution (-5.5 kJ/mol). The five H-bonds formed by the carboxylate group and the nitro function of **30** with several residues of ALR2 contribute favorably to the enthalpy. As **30** and **31** show a very similar binding mode also their thermodynamic profiles are comparable. **31** has a slightly less favorable enthalpic portion compared to **30** (-24.4 kJ/mol compared to -27.4 kJ/mol) which is compensated for by a small more favorable entropic contribution (-8.2 kJ/mol compared to -5.5 kJ/mol). The only difference between both ligands is the exchange of the terminal phenyl ring of **30** by a furan ring in the case of **31**. As discussed above the oxygen atom of the furan ring is not involved in any further H-bonds. Thus the loss of its solvation shell will contribute unfavorably to the enthalpic binding portion since the H-bond inventory experienced in solution with surrounding water molecules is unbalanced. On the contrary the released water molecules from the solvation shell of **31** will increase the degrees of freedom of the overall system and explains, at least in part, why **31** possesses a more favorable entropic portion compared to **30**. Furthermore the π-π-stacking should be less efficient for **31** compared to **30**. The furan ring of **31**

has due to its oxygen atom an enhanced electron density compared to the phenyl ring of **30**. Since Trp111 possesses an electron-rich indole ring an electron-deficient system will be preferred as interaction partner, also indicated by the fact that electron-withdrawing substituents of this ring such as nitro-groups, fluorine and chlorine are beneficial for binding.[43] Electron deficiency is better fulfilled for **30** as for **31** which could additionally explain the slight enthalpic advantage of **30** compared to **31**. **26** shows a very similar thermodynamic profile as **30**. The enthalpic portion (-27.2 kJ/mol compared to -27.4 kJ/mol) and the entropic portion (-4.3 kJ/mol compared -5.5 kJ/mol) are nearly identical despite remarkable changes in the binding modes. Steuber et al.[8] discussed several factors which influence the thermodynamic profile of **26**. For the relative discussion of **26** and **30** these assumptions were taken into account. Several factors influence the enthalpic contribution of the binding of **26** to ALR2. The formed H-bonds between **26** and the captured water W1 and between W1 and Trp111 are contributing favorably to the enthalpic portion. In ALR2-**30** no similar interaction to a water molecule can be detected. Accordingly **26** should have an enthalpic advantage compared to **30**. A small enthalpic advantage arises from the H-bond between one of the oxygen atoms of the carboxylic acid group of **26** and the Tyr48OH. It is in ALR2-**26** shorter in comparison to the other complexes, possibly due to the fact that this oxygen atom is in contrast to the other complexes not involved in any further H-bonds. Therefore, it can be assumed that it is accompanied by a stronger favorable enthalpic contribution compared to the comparable contact in ALR2-**30**. On the other hand, some characteristics of the chemical structure and the binding mode of **26** influence the enthalpy unfavorably. In all other complexes except ALR2-**28**-$P2_1$ and ALR2-**28**-P1 an H-bond (2.9-3.1 Å) between one of the oxygen atoms of the carboxylate group of the inhibitor and Trp111NE1 is observed. Indeed these distances indicate a rather weak interaction. Nevertheless, this missing interaction in ALR2-**26** should be accompanied by a loss in the enthalpy and has to be overcompensated by the interactions involving W1. Furthermore, the central oxadiazol ring of **26** will be well solvated in solution. The situation can be compared to that of the furan ring in **31**. An unfavorable effect on the enthalpy can be assumed since the heteroatoms of the oxadiazol moiety of **26** are not compensated for the lost interactions to the solvent water molecules by polar contacts to the protein residues. Regarding the entropy the following assumptions should be regarded. The degrees of freedom of water W1 are decreased which contributes unfavorable to the entropic

profile. On the other hand the water molecules which are part of the solvation shell of **26** are released when **26** binds to the protein. This in consequence leads to a favorable entropic contribution and compensates, at least in parts, for the unfavorable entropy loss accompanied with the capturing of the water molecule W1. Altogether the comparison of **26** and **30** is an example where significant differences in the binding mode finally display a very similar thermodynamic signature. **28** shows a more favorable enthalpic (-33.5 kJ/mol) and a less favorable entropic portion (-1.3 kJ/mol) than **30**. This is in line with the deviating binding modes. In ALR2-**28**-$P2_1$ two water molecules (W2 and W3) are captured. Again, as discussed for **26**, the formed H-bonds between **28**, W2, W3 and the protein residue will result in a gain in enthalpy compared to **30** whereas the entropic part is reduced owing to the fixation of the water molecules. The established H-bond network of **28** involving W2 and W3 leads to such a large gain in enthalpy most likely overcompensating the missing interaction between the carboxylate group of **28** and Trp111. Both effects, the gain in the enthalpic contribution and the loss in the entropic contribution are more distinct for **28** than for **26** since **28** captures two waters instead of one. Surprisingly, **28** and **29** deviate strongly in their thermodynamic profiles. The enthalpic portion is -33.5 kJ/mol for **28** while it is -23.4 kJ/mol for **29**. The entropic portion is -1.3 kJ/mol for **28**, but it amounts to -8.9 kJ/mol for **29**. Despite several trials we were not able to obtain the complex structure of **29**. In consequence we can only hypothesize about a possible explanation of its deviating thermodynamic signature. It was previously discussed that the carboxylic acid group of **29** might adopt a different orientation compared to the one of **28**.[143] However, the comparison of ALR2-**28**-$P2_1$ and ALR2-**30** indicates moreover that the different central aromatic ring systems lead to deviating interactions of the carboxylic acid functions of both ligands. As shown in Figure 4-5e-f, the relative positions of the C2-linker between the respective aromatic ring system and carboxylic acid function differ strongly since a phenyl and thiophenyl ring do not allow for the attachment of the C2 linker with the same geometry. Consequently, only the α-oxygen atoms of both ligands align while the β-oxygen atoms differ in their positions. Since **29** contains similar to **28** a five-membered central aromatic moiety the binding geometries of the carboxylic acid group should be similar for both ligands. Although no explanation for the deviating thermodynamic profiles of **29** and **28** can be provided this example demonstrates how

significantly the exchange of a thiophene ring to a furan ring can influence the thermodynamic profile.

## 4.5 Conclusions

Due to its role in the polyol pathway ALR2 is a promising target to prevent the development of diabetic complications. Here, a series of closely related carboxylic acid type ligands bearing two or three aromatic ring moieties were evaluated as inhibitors of ALR2. It could be confirmed that the introduction of the *meta* nitro group to the terminal aromatic ring system is accompanied by a potency enhancement which is in concordance with previous studies. Surprisingly, the exchange of this nitro group by a carboxylic acid group leads to a significantly decreased affinity. **37** is with an $IC_{50}$-value of 3.1 +/- 0.8 µM an interesting derivative bearing an acrylic acid group while **43** is with an $IC_{50}$-value of 0.34 +/- 0.05 µM[144] a potent member of the compound class containing a third aromatic ring system. The soaking protocol was improved so that the pH value facilitates the binding of the inhibitors. Solving the complex crystal structure of **28** complexed to ALR2 in space group $P2_1$ (ALR2-**28**-$P2_1$) showed that the binding mode of **28** is in good agreement with the previously described corresponding structure in P1. Furthermore, an additionally captured water molecule (W3) is visible in the ALR2-**28**-$P2_1$ complex which might influence the thermodynamic profile. The binding mode of four new ALR2-inhibitor complexes (ALR2-**30**, ALR2-**31**, ALR2-**37**, ALR2-**43**) were determined. ALR2-**37** is the first complex structure of an ALR2 inhibitor complex which contains an acrylic acid head group. Since the conformation of the double bond of **37** changed from *E*-to *Z*-configuration prior to binding it can be deduced that ALR2 prefers inhibitors with the acrylic acid group in *Z*-configuration. ALR2-**43** is the first complex structure of ALR2 with an inhibitor of this compound class which exhibits three aromatic moieties. The structure shows that ALR2 can accommodate such a bulky ligand.

Additionally, the thermodynamic profiles could be measured in HEPES buffer. The enthalpic binding portion across the studied ALR2- inhibitor series varies between the lowest and largest value over a range of 10.1 kJ/mol. This is remarkably large since the binding modes of the ligands to the target protein ALR2 are closely comparable. In two cases captured water molecules could explain the differences in the thermodynamic profiles. Different interaction patterns of the carboxylic acid head group have been observed. In some cases the burial of polar groups of the

inhibitors in a hydrophobic environment of the protein provide an explanation for a reduced enthalpic signature. The relationship between the thermodynamic binding patterns and the adopted binding modes described here can serves as a valuable input for further design cycles.

## 4.6 Description of Experimental Procedures

### 4.6.1 Expression, purification, ITC measurements, crystallization and structure determination

The expression and purification of ALR2 are described in Section 9.2.1, the ITC measurements in Section 9.2.5, the crystallization in Section 9.2.3, the soaking in Section 9.2.3.6 and 9.2.3.7 and the structure determination in Section 9.2.4.

### 4.6.2 $^{1}$H-NMR analysis of 37

No sign of the *Z*-isomer of **37** was visible within the compound´s $^{1}$H-NMR spectrum when it was measured immediately after the synthesis. The doublet at 6.69 ppm with a coupling constant of 16.0 Hz clearly shows the presence of the *E*-isomer. However, a doublet at 6.02 ppm with a coupling constant of 12.6 Hz and an integral of 0.02 (integral of the *E*-isomer doublet = 1.00) appeared in the spectrum after three years. This signal with its typical coupling constant belongs to the *Z*-isomer.[150] Therefore we concluded that an isomerization of the double bond had taken place during storage. The compound was stored undissolved in a box, but no additional measures were taken to protect it from light. An attempt to separate the isomers chromatographically (silica gel and dichloromethane / methanol mixtures) was unsuccessful. Instead the *Z*-isomer was enriched. The NMR spectrum showed the *Z*-isomer doublet at 6.02 ppm with an integral of 0.15 (integral of the *E*-isomer doublet = 1.00). Whether the enrichment of the *Z*-isomer was only due to isomerization under the conditions of chromatography or whether a more favorable elution of the *Z*-isomer played a part could not be determined. The NMR sample of the mixture (dissolved in deuterated DMSO) was exposed to ultraviolet light (254 nm) and subsequently spectra were taken. After 8 h exposure the spectrum showed the *Z*-isomer doublet at 6.02 ppm with an integral of 0.39.The *E*-isomer doublet at 6.69 ppm was superimposed with the second doublet of the *Z*-isomer in a way that made separate integration impossible. The combined signal was set to 1.00. This means the ratio of *E* to *Z* was around 3:2 at this time and continued to increase during exposure to ultraviolet light. The doublet of the *Z*-isomer was integrated to 0.44 after 16 h

exposure and to 0.52 after 31 h (signal at 6.69 ppm was always set to 1.00). Therefore it seems clear that the geometry of the double bond is liable to isomerization and especially unstable in solution, under exposure to ultraviolet light as well as during chromatography while even the undissolved solid is not completely stable.

# 5 Thermodynamic and Structural Investigation of the Opening of the Specificity Pocket of Human Aldose Reductase

## 5.1 Introductory Remarks

The following chapter was prepared for submission to a scientific journal. This study was done in collaboration with Dr. Frithjof Scheer, Felix Terwesten, Dr. Philipp Toth, Prof. Dr. Andreas Heine and Prof. Dr. Wibke Diederich (Philipps University Marburg). The author of this thesis determined the complex crystal structures, performed the ITC measurements and drafted the following manuscript.

## 5.2 Abstract

For drug design a profound understanding of the process of molecular recognition between the binding site of a protein and small molecules which have the ability to inhibit this protein is essential. Here, we present insights into the thermodynamic background of the opening of a subpocket of a protein when addressed by a suitable moiety of a ligand. As a model protein the human aldose reductase was chosen, which is involved in the development of diabetic complications. The opening of its specificity pocket, which is highly significant for the design of selective inhibitors for this protein, is studied by high-resolved X-ray crystallography, isothermal microcalorimetry and surface plasmon resonance measurements, using a ligand series based on a 2-carbamoyl-phenoxy-acetic acid scaffold.

## 5.3 Introduction

For drug design a detailed understanding of the thermodynamic behavior of the binding process of a ligand to a protein is essential.[6, 7] This appears even more important if for the accommodation of the ligand a transient subpocket has to be opened. Little is known about the influence of the opening process of such a subpocket on the thermodynamic profile. To gain further insights into this process the choice of human aldose reductase (ALR2, EC 1.1.1.21) as a model protein appears promising. Due to its involvment in the development of diabetic complications this protein has medicinal relevance. In 2013, 382 million people worldwide suffered from diabetes. This number is predicted to increase to 592 million cases up to the year 2035.[17] The disease is related to some long-term complications like neuropathy, nephropathy

and retinopathy.[45] ALR2 is a NADPH-dependent oxidoreductase. It reduces glucose to sorbitol which is the first step of the polyol pathway. The consumption of NADPH leads to decreased levels of reduced glutathione[21] which makes the cells more vulnerable for intracellular oxidative stress. This is a confirmed mechanism which leads to the development of diabetic complications,[28, 29] making ALR2 an interesting target for drug design. This made ALR2 to a well characterized protein. Ultrahigh resolved crystal structures[40] and thermodynamic data of protein ligand complexes, giving insights into key interactions[8] and overlaid protonation effect[42] are described in literature. This knowledge served as a starting point for the current study.

The binding pocket of ALR2 is depicted in Figure 5-1a. One subpocket consists of an anion binding pocket which is formed by Tyr48, His110, Trp111 and the nicotinamide moiety of $NADP^+$.[38] Inhibitors address this subpocket with a negatively charged group, e.g. their carboxylate, hydantoin or succinimide functions.[45] Adjacent, an additional subpocket can be formed in the active site. The C-terminal loop, especially the region around Leu300, possesses a high flexibility, which allows the opening of the so-called specificity pocket. Phe122 and Leu300 serve as gatekeepers which either seal or give access to this pocket.[37] In Figure 5-1b the binding mode of fidarestat (**44**, PDB: 1PWM)[151] is shown. It does not address the specificity pocket which remains in the closed state. In Figure 5-1d the binding mode of minalrestat (**45**, PDB: 1PWL)[151] is depicted. The latter compound addresses the specificity pocket with its halogenated phenyl substituent. The side chain of Leu300 (colored in red) shifts its position in space and opens the entrance of the pocket. In Figure 5-1c the different conformations of Phe122 and the Ala299 to Ser302 sequence stretch in both structures are compared. The image shows that not only the side chain of Leu300 but the entire attached protein backbone changes its conformation. Accommodation of ligand portions in this pocket by inhibitors which exhibit suitable moieties is a promising strategy to gain specificity compared to the human aldehyde reductase (EC 1.1.1.2). For the latter enzyme the formation of such a pocket is less favorable since it would involve the rupture of a salt bridge.[39] The crystal structure of the ALR2 *holo*-enzyme shows the closed conformation of the specificity pocket (PDB: 1ADS).[33] In the 138 complex crystal structures which are deposited in the PDB (August 2015) the specificity pocket

is only opened if addressed by an inhibitor. Consequently, the closed conformation appears to be the thermodynamically more stable one.

A previously published class of ALR2 inhibitors possess 2-arylcarbamoyl-phenoxy-acetic acid scaffold as the core structure.[43] In the present work, we use derivatives of this group with varying moieties (R in Figure 5-2) to induce different binding site conformers. The chemical structures of the studied inhibitors are presented in Figure 5-2. The resolution of the here described complex crystal structures ranges from 0.98- 1.20 Å which allows for a detailed analysis of the binding modes of the inhibitors, the induced conformation of the binding pocket and water structures. Isothermal microcalorimetry (ITC) and surface plasmon resonance (SPR) measurements were used to study the thermodynamic background of the addressing and consequently opening of the specificity pocket.

## 5.4 Results

### 5.4.1 Introductory remarks regarding the crystallization and soaking conditions

The complex structures ALR2-**47**, ALR2-**49**, ALR2-**50**, ALR2-**52** and ALR2-**53** were obtained by soaking the ligands in TRIS buffer at pH 8.0 into *holo*-protein crystals. ALR2-**51** was obtained by soaking in citrate buffer at pH 5.0. ALR2-**48** was obtained by cocrystallization in citrate buffer at pH 5.0 since this ligand destroyed the crystals during the soaking process in TRIS buffer. Since the differences which we found in this study between the binding modes of the different ligands can be explained by conformational adaptions on the protein side or by displacement of water molecules due to sterical reasons we consider the influence of the different crystallization and soaking conditions as negligible.

### 5.4.2 Different protein conformers

Three different conformers of the residues which border the specificity pocket were found in the here described crystal structures. Figure 5-1e-j, Figure 5-4 and Figure 5-7 give an overview about the induced protein conformers of ALR2-**47**- ALR2-**52**. First, ALR2-**47**, ALR2-**48** and ALR2-**52** represent complex crystal structures where the specificity pocket is found in its closed state. Here, the side chain of Leu300 points towards Phe122 which blocks the access to the pocket.

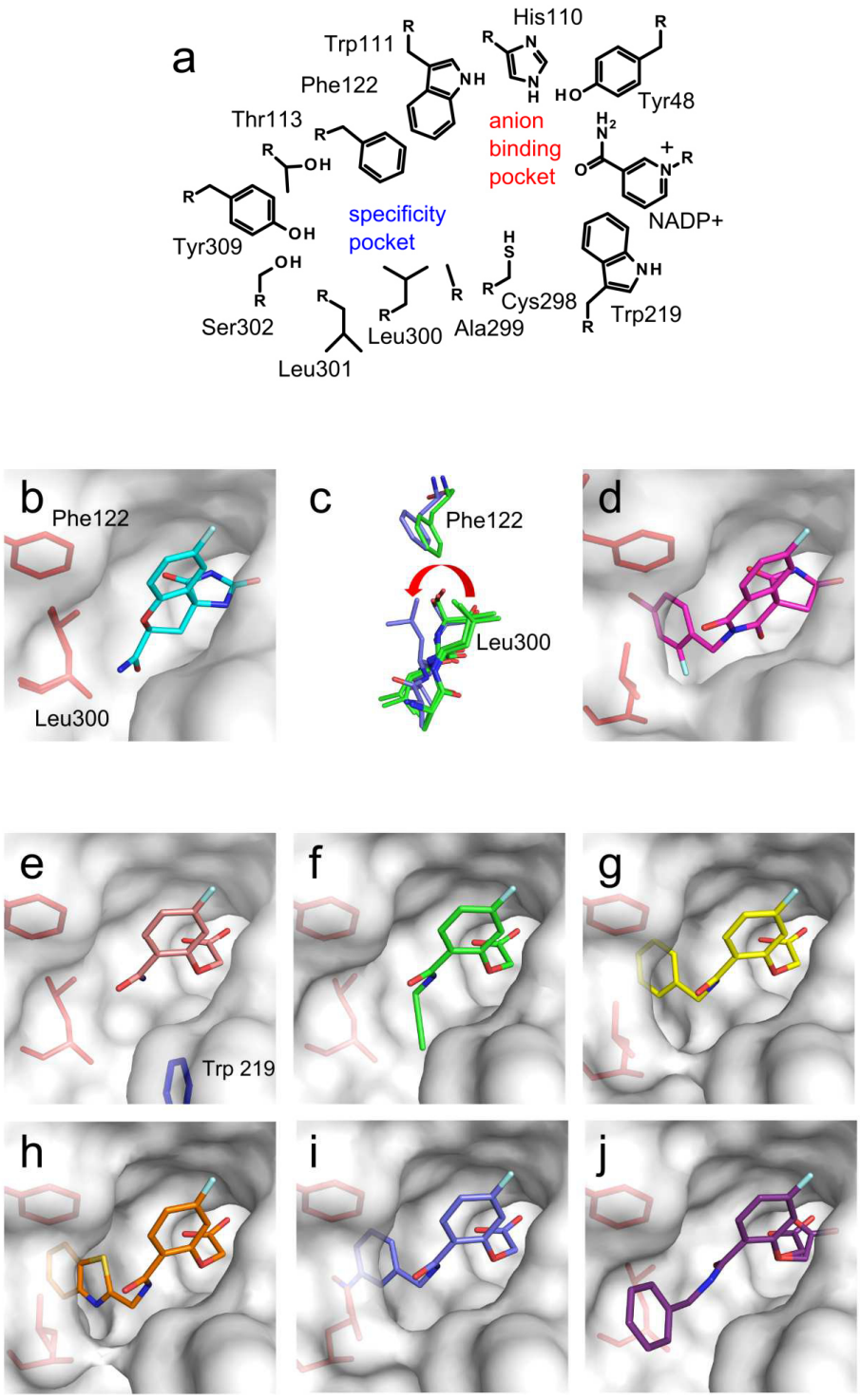



**Figure 5-1** Detailed figure caption on next page.

**Figure 5-1** (a) Schematic view of the binding pocket of ALR2. (b) Binding mode of **44** with the specificity pocket in closed conformation. The protein surface is colored in gray, the carbon atoms of **44** are colored in cyan, the oxygen atoms in red and the nitrogen atoms in blue. Phe122 and Leu300 are colored in red. This color scheme is maintained during the following images. (c) The view direction is perpendicular compared to the remaining figures. The carbon atoms of the protein residues which correspond to ALR2-**44** (specificity pocket is closed) are colored in green, those which correspond to ALR2-**45** (specificity pocket is opened) are colored in blue. The entire backbone next to Leu300 performs conformational changes when the specificity pocket is opened. This is best visible when the positions of the side chain of Leu300 are compared (red arrow). (d) Binding mode of **45** with the specificity pocket in open conformation. Carbon atoms of **45** are shown in purple. (e)-(j) Binding modes of the ligands which are investigated in this study. **47**, **48** and **52** leave the specificity pocket closed while **49**, **50** and **51** open the specificity pocket. (e) ALR2-**47**, carbon atoms of **47** are shown in light-red. (f) ALR2-**48**, carbon atoms of **48** are shown in green. (g) ALR2-**49**, carbon atoms of **49** are shown in yellow. (h) ALR2-**50**, carbon atoms of **50** are colored in orange. (i) ALR2-**51**, carbon atoms of **51** are colored in blue. (j) ALR2-**52**, carbon atoms of **52** are colored in purple. This figure and the following representations were prepared using PyMOL.[36]

As a consequence of the position of Leu300 the protein surface forms a cleft between the backbone of Leu300 and the side chain of Trp219 (Figure 5-1e, f and j). Secondly, ALR2-**49** and ALR2-**51** exhibit the specificity pocket in an open conformation. The loop around Leu300 performs conformational changes to open the specificity pocket. The side chain of Leu300 moves away from the phenyl ring of Phe122. Leu300, Cys298 and Trp219 form a kind of wall which limits the binding pocket at the lower side (Figure 5-1g, i). **53** induces with the *meta* nitro phenyl moiety the same conformation of the residues of the specificity pocket as **49** and **51** (not shown in Figure 5-1). Thirdly, in ALR2-**50** the specificity pocket is similarly opened, but here the carboxamide group of the Ala299- Leu300 backbone is orientated differently (Figure 5-3a). An H-bond between the nitrogen of the benzothiazole ring and the nitrogen of the Ala299-Leu300 amide bond leads to this rearrangement of the backbone of the specificity pocket (Figure 5-4d). Figure 5-3b shows an alignment of Ala299-Ser302 of the obtained complex crystal structures. The conformational changes result in the largest repositioning of the side chains of Leu300 and Leu301. The transformation from the closed to the opened state of the pocket is accompanied by the rupture of one and the formation of two new internal H-bonds between the residues which border the specificity pocket. In Figure 5-3c an overlay of ALR2-**47** (carbon atoms are colored in salmon, specificity pocket is in closed state) and ALR2-**51** (carbon atoms are colored in blue, specificity pocket is in open state) is presented. In the closed state, an H-bond between the oxygen atom of the carbonyl group of the peptide bond between Ala299 and Leu300 and the Tyr309OH group is established. The short distance of 2.5 Å indicates a strong interaction. In the open state the peptidic carbonyl oxygen is turned away from Tyr309 (Figure 5-3d).

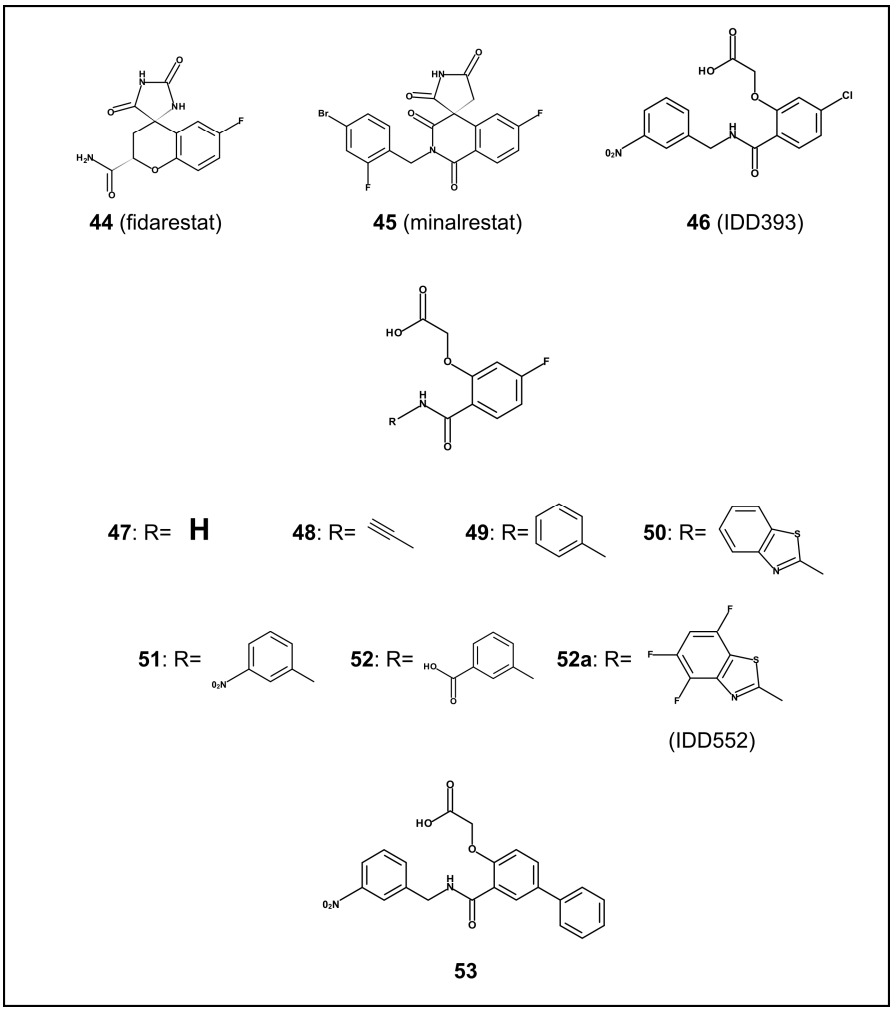



**Figure 5-2 Chemical formulas of the studied inhibitors**

Thus, this H-bond cannot be formed any longer. In contrast, two new but longer H-bonds are formed, involving the oxygen atom of the amide carbonyl oxygen of the Leu301-Ser302 backbone: one is established with the nitrogen of the amide bond of the Cys303-Thr304 backbone (3.2) and one with the OH group of Thr304 (2.9 Å, Figure 5-3d). The latter H-bonds are also formed in the slightly different conformation of the specificity pocket in ALR2-**50** (not shown in Figure 5-3d). Consequently, the analysis of the internal H-bonds of all the three here detected conformations of the residues of the specificity pocket indicates that the newly formed internal H-bonds in the open conformation compensate for the loss of the internal H-bond

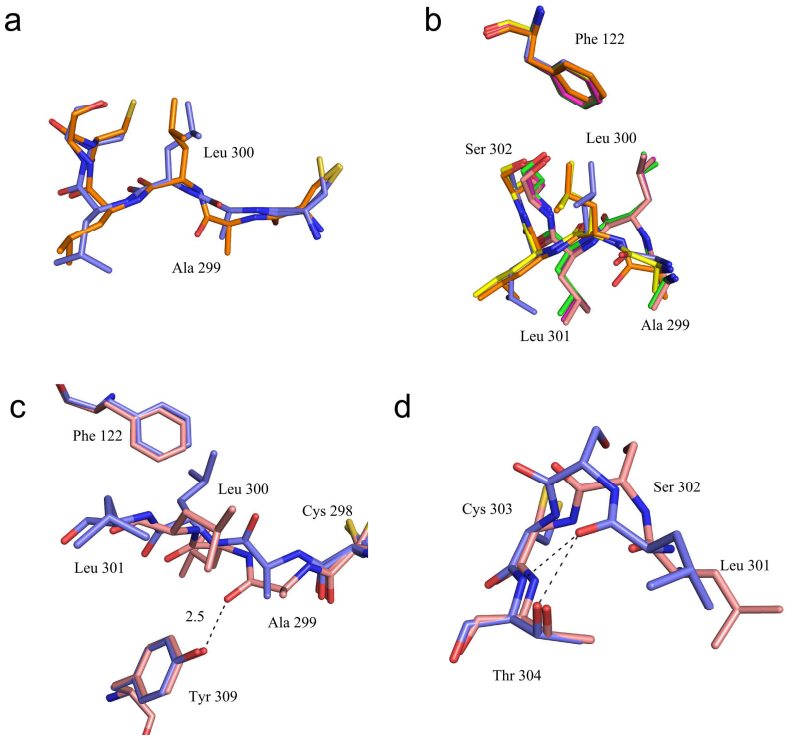



**Figure 5-3 (a)** Comparison of Cys298-Cys303 of ALR2-**51** (carbon atoms are shown in blue) and ALR2-**50** (carbon atoms are shown in orange). (b) Alignment of Phe122 and Ala299 to Ser302 of ALR2-**47** (carbon atoms are shown in light red), ALR2-**48** (carbon atoms are shown in green), ALR2-**49** (carbon atoms are shown in yellow), ALR2-**50** (carbon atoms are shown in orange), ALR2-**51** (carbon atoms are shown in blue) and ALR2-**52** (carbon atoms are shown in purple). (c) Alignment of ALR2-**47** and ALR2-**51**. An internal H-bond between the residues Ala299 and Tyr309 in the closed conformation of the specificity pocket is drawn as a black-dotted line. Carbon atoms of the protein residues of ALR2-**47** which represent the closed conformation of the specificity pocket are shown in light red, the carbon atoms of the corresponding atoms of ALR2-**51** which represent the open conformation are shown in blue. (d) Alignment of ALR2-**47** and ALR2-**51**. Same color scheme as in (c) was applied. Two internal H-bonds between the residues Ser302 and Thr304 are formed in the open conformation of the specificity pocket.

observed in the closed conformation. However, the different adopted geometries might still involve a change in the energy inventory of the residues in the different states.

### 5.4.3 Binding mode of the 2-(2-carbamoyl-5-fluoro-phenoxy)acetic acid scaffold

**47**-**52** all possess 2-(2-carbamoyl-5-fluoro-phenoxy)acetic acid as parent scaffold, which performs identical interactions in all cases with the enzyme except for the diacid **52** (Figure 5-4, Figure 5-7). The carboxylic acid function of the scaffold is directed into the anion binding site.

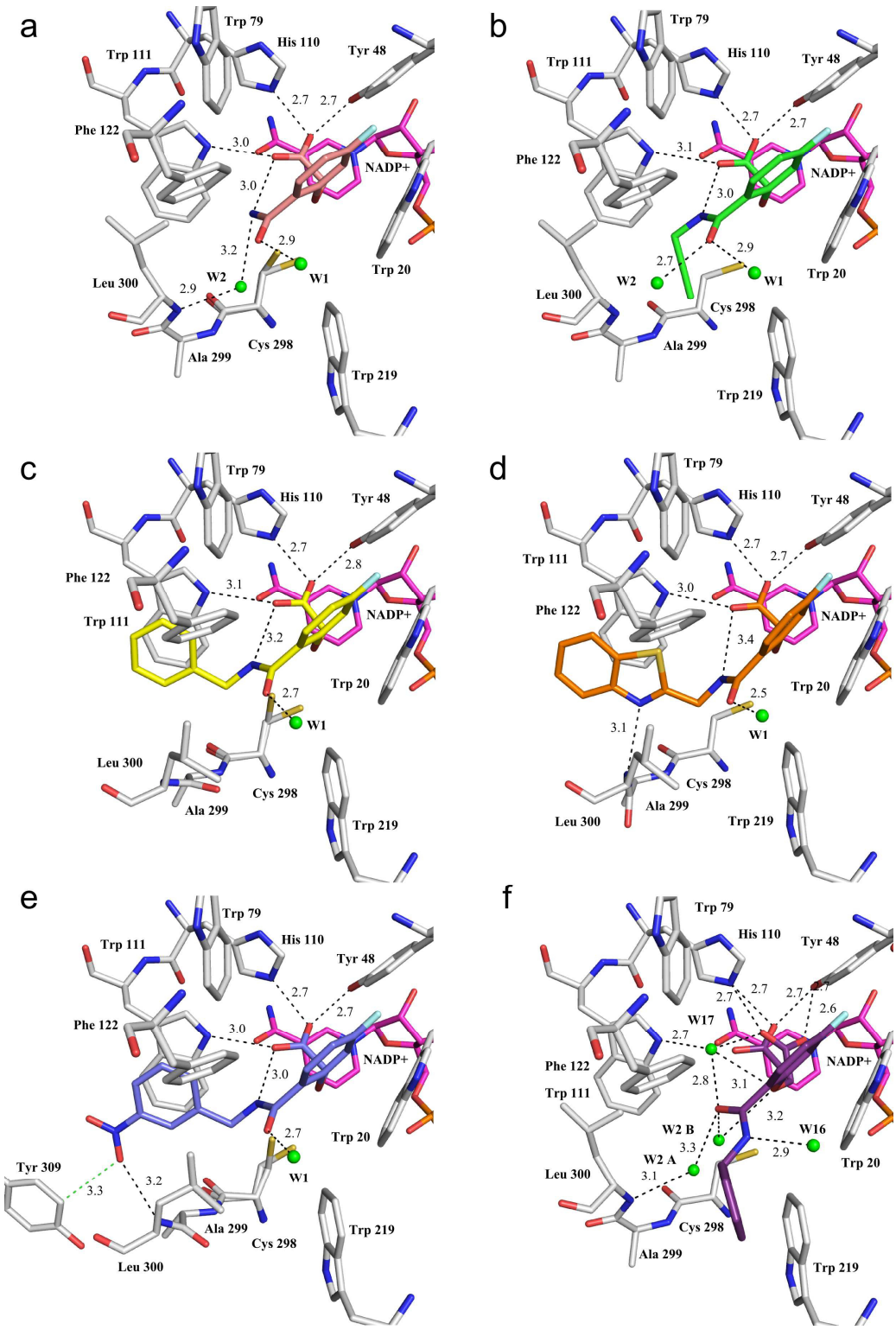



**Figure 5-4** Detailed figure caption on next page.

**Figure 5-4 Stick representation of the binding modes of 47-52.** Carbon atoms are colored in gray, nitrogen atoms are colored in blue, sulfur atoms in yellow and oxygen atoms in red for protein residues. Carbon atoms of $NADP^+$ are colored in magenta. Water molecules which are interacting with the ligand are shown as green spheres, H-bonds are depicted as black-dotted lines. (a) ALR2-**47**, carbon atoms are shown in light red. (b) ALR2-**48**, carbon atoms are shown in green. (c) ALR2-**49**, carbon atoms are shown in yellow. (d) ALR2-**50**, carbon atoms are shown in orange. (e) ALR2-**51**, carbon atoms are shown in blue. The non-classical H-bond between one of the oxygen atoms of the nitro function of **51** and one of the hydrogen atoms of the phenyl ring of Tyr309 is displayed as a green-dotted line. (f) ALR2-**52**, carbon atoms are shown in purple.

In a previous study the cofactor was described as being bound in its oxidized form $NADP^+$.[40] Figure 5-4 shows stick representations of ALR2-**47** to ALR2-**52** along with the polar interactions formed between ligand, protein residues and water molecules, Figure 5-7 the same complexes with corresponding water networks. The complex structures were obtained at pH 8.0 and pH 5.0, respectively. At pH 8.0 the carboxylate group should be deprotonated during the entire binding process. At pH 5.0 it should be deprotonated in bound state, however prior to protein binding it should be partly protonated. Therefore it will carry a negative charge on the acetic acid group by which it interacts with the positively charged nicotinamide moiety of $NADP^+$. The distance between the carboxylate oxygen which falls near to the cofactor and the endocyclic nitrogen of the nicotinamide of $NADP^+$ is 4.2 Å. Furthermore, the acetic acid carboxylate oxygen forms H-bonds to Tyr48OH (2.7-2.8 Å) and His110NE2 (2.7Å) while the other oxygen of the carboxylate group establishes an H-bond to Trp219N1 (3.0-3.1 Å). The central peptide bonds of **47**-**51** position in a way that an internal H-bond between the amide nitrogen NH to one of the oxygen atoms of the carboxylate group of the inhibitor (3.0-3.4 Å) can be formed. Solely **52** does not form this interaction. All investigated ligands (except **52**) form an H-bond contact (2.5-2.9 Å) between the carbonyl oxygen of the central peptide bond and water W1. Additionally the central fluoro-substituted phenyl ring makes van der Waals contacts to Trp79 (4.6 Å), Trp20 (3.8 Å) and Phe122 (3.9 Å). In the following Section the differences between the binding modes of ligands **47**-**53** are explained in detail.

### 5.4.4 Binding mode of 47

Since **47** carries no further substituent at the amide nitrogen the doubly protonated nitrogen can undergo an additional H-bond contact to the additionally recruited water molecule W2 (3.2 Å, Figure 5-4 a) which in turn is H-bonded via an additional water molecule W3 with the nitrogen of the amide bond between Ala299/Leu300 (2.9 Å). Consequently, using W2 and W3 as mediating

water molecules, **47** forms additional polar contacts to the protein. In the here investigated ligand series it is the only ligand besides **52** which forms an H-bond via this amide nitrogen in addition to the intramolecular H-bond to the terminal acetic acid carboxylate group.

### 5.4.5 Binding mode of 48

**48** contains a propynyl side chain which is attached to the amide function. The propynyl side chain contains a π-electron system and fills the cleft next to the unopened specificity pocket formed between Leu300 and Trp219 which has a rather hydrophobic character. This binding mode enables the propynyl side chain of **48** to form van der Waals contacts to Trp111 (3.4 Å), Phe122 (4.8 Å), Trp219 (3.6 Å) and Leu300 (3.5 Å). To accommodate the side chain of **48** Ala299–Tyr309 have to shift slightly compared to the ALR2-**47** structure (Figure 5-5a). Using the McLachlan algorithm[152] as implemented in the program ProFit[153] an RMSD across the residues 299-309 of 0.67 Å was calculated which is 3.6 times larger than the RMSD of the remaining protein residues (0.19 Å) in this comparison. For this calculation and the additional calculations referenced further down with ProFit the protein molecules were aligned based on the $C_\alpha$ atoms. Similar to **47** also **48** forms a direct H-bond to W2 (2.7 Å), however, here the ligand uses its carbonyl oxygen of the central amide bond to establish this contact (Figure 5-4b).

### 5.4.6 Binding mode of 49

**49** could be built into the $F_o$-$F_c$ difference electron density map at a σ level of 3.0. It was refined to an occupancy of 80%. Still unassigned significant difference electron density remained in the specificity pocket. An overlay with ALR2-**47**, in which the specificity pocket keeps the closed conformation, shows a univocal match between the backbone and side chain positions of the residues of the specificity pocket and the unexplained difference electron density (Figure 5-5b). This observation strongly supports the presence for a second conformation of the specificity pocket in ALR2-**49**. Predominantly the open conformation is visible which is induced by **49** binding. It was adjusted to the same occupany as the ligand (80%). The protein residues which correspond to the open conformation of the specificity pocket adopt a similar geometry as in ALR2-**51**. Because the corresponding $F_o$-$F_c$ electron density is not fully visible we restrained from inserting the closed conformation of the specificity pocket as a second conformation. The question remains whether **49** binds in two conformations, in agreement with the observed two

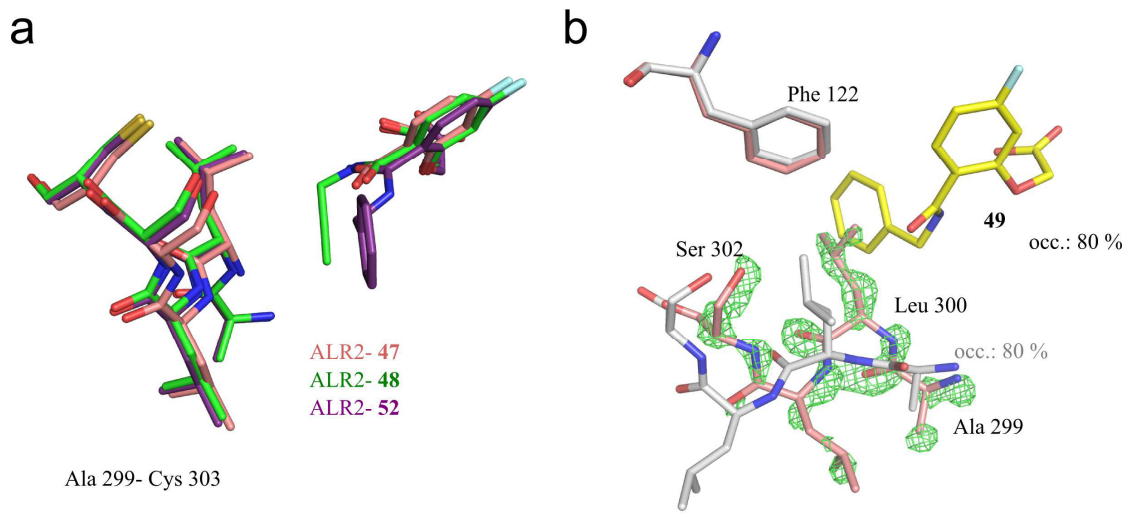



**Figure 5-5** (a) Alignment of ALR2-**47**, ALR2-**48** and ALR2-**52**. The carbon atoms of protein residues Ala299 to Cys303 and of the ligands are shown in light red for ALR2-**47**, in green for ALR2-**48** and in purple for ALR2-**52**. In ALR2-**48** and ALR2-**52** the residues Ala299 to Cys303 are pushed aside to accommodate the bulky moieties of the corresponding ligand. (b) **49** is shown together with Ala299- Ser302. The residues with carbon atoms shown in gray correspond to the open conformation of the specificity pocket which was refined to an occupancy of 80%. This conformation was included in the final model. Some $F_o$-$F_c$ difference electron density was found at a level of 3.0 sigma which corresponds to the closed conformation of the specificity pocket (carbon atoms are shown in light red). This can be attributed to 20% of the protein molecules where the specificity pocket is not opened by a molecule of **49**.

conformers of the specificity pocket. In such case, the difference electron density for a second ligand conformer with respect to the terminal phenyl ring of **49** should be visible next to the entrance of the binding pocket. However, we did not find any $F_o$-$F_c$ difference electron density suggesting the occupancy of the phenyl ring in this area at a σ level of 3.0. Additionally, the average *B*-factor of the terminal phenyl ring of **49** (13.1 $Å^2$) is only slightly higher than that of the attached fluorophenoxyacetic acid scaffold (10.4 $Å^2$) which binds tightly to the anion binding pocket. If a second orientation of the terminal phenyl moiety of the ligand existed the terminal phenyl ring should show much higher *B*-factors. Consequently, the averaged *B*-factors suggest that supposedly no second orientation of the ligand is present and that **49** only binds if the specificity pocket is opened. There, the terminal phenyl ring of **49** accommodates the specificity pocket and performs a π-π-stacking with Trp111 (~3.5 Å). Furthermore, the phenyl ring undergoes van der Waals contacts to the aromatic systems of Trp79 (3.7 Å) and Phe122 (3.7 Å) and to the side chain of Leu300 (3.6 Å). Up to now this is the only known complex crystal

structure which shows the binding of an unsubstituted aromatic system in the specificity pocket of ALR2. With a $K_D$= 7.5+/-1.4 µM it is one of the weakest binders, confirmed by X-ray crystallography, to open the specificity pocket of ALR2 (Figure 5-4c).

### 5.4.7 Binding mode of 50

An H-bond is formed between the amide nitrogen of the Ala299-Leu300 bond and the nitrogen atom of the benzothiazolyl moiety of **50** (3.1 Å, Figure 5-4d). Therefore, the backbone changes its position to enable the formation of this H-bond compared to ALR2-**51** (Figure 5-3a). The benzothiazolyl moiety of **50** forms a face to face π-π-stacking with the side chain of Trp111 (approximately 3.5 Å). Furthermore, it forms van der Waals contacts to Trp79 (4.2 Å), Phe122 (4.0 Å) and Leu300 (3.4 Å). The sulfur atom of the benzothiazolyl moiety is not involved in any polar interactions.

### 5.4.8 Binding mode of 51

**51** possesses a nitro group attached in *meta* position to its terminal phenyl ring. The latter moiety of **51** opens the specificity pocket (Figure 5-4 e and Figure 5-7e). Its binding mode is very similar to that of its Cl-analogue IDD393 (**46**).[8] The phenyl ring of the side chain performs a π-π-stacking with the aromatic system of Trp111 (~3.5 Å). Additionally, it forms van der Waals contacts to the side chains of Trp79 (3.8 Å) and Phe122 (3.8 Å). Thus, the hydrophobic interactions of the phenyl ring of **51** are similar as already described for the corresponding moiety of **49**. The nitro group forms an H-bond to the NH of the Leu300 backbone (3.2 Å). Furthermore, one of the oxygen atoms of the nitro group is positioned close to one of the carbon atoms of the phenyl ring of Tyr309 (3.3 Å, Figure 5-4 e). This distance is surprisingly short. It was previously discussed to be a non-classical secondary H-bond between the C-H dipole at the phenyl ring and the negatively polarized oxygen of the nitro group.[8]

### 5.4.9 Binding mode of 52

The occupancy of **52** was refined to 77%. Compared to **51** the nitro function is replaced in **52** by an isosteric carboxylic acid function. Unfortunately, the terminal carboxylic acid function at the phenyl ring is not visible in the $F_o$-$F_c$ difference electron density at a σ level of 3.0 and was consequently not included in the deposited structure. The NMR, mass-spectrometric and

elemental analysis of **52** show unambiguously the presence of the latter function (see Appendix 10.3). We explain the missing electron density with an increased residual mobility of the benzoic acid moiety which is oriented toward the solvent-accessible area of the binding pocket. The acid group cannot form any H-bonds to protein residues which could assist to fix the group it in a certain position. The high mobility is reflected by an increased mean *B*-factor of the phenyl ring of the benzoic acid moiety (14.5 $Å^2$) compared to the carbamoyl-fluoro-phenoxyacetic acid scaffold (7.1 $Å^2$). To confirm the result the experiment was repeated independently from the first trial. The same soaking protocol was applied. In the second diffraction experiment even less $F_o$-$F_c$ difference electron density for the benzoic acid moiety of **52** was visible at a σ level of 3.0 (Figure 5-6).

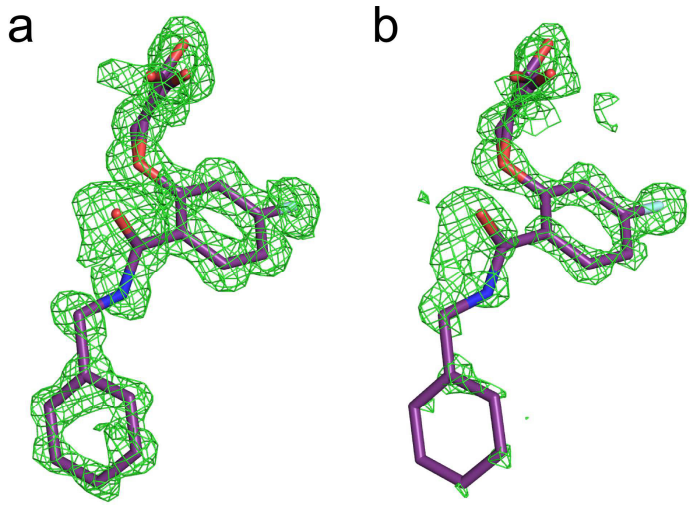



**Figure 5-6 Comparison of the $F_o$-$F_c$ difference electron densities (shown as a green mesh) of 52 in complex with ALR2 at a σ level of 3.0**: (a) dataset which was deposited and (b) dataset from a second diffraction experiment with another crystal obtained by the same soaking condition. (a) The terminal *meta* carboxylate group is not visible in the density but the remaining molecule is well defined. (b) The terminal phenyl ring is only partly visible in the density indicating even higher mobility of this moiety than of the remaining part of the molecule.

This result confirms the assumption that the benzoic acid moiety of **52** has a higher mobility than the remaining part of this ligand. The high-energetic synchrotron radiation could principly lead to radiation damage, resulting in bond cleavage between the carboxylic acid function and the terminal phenyl ring. Prior to the dataset collected at the synchrotron (BESSY II, Berlin-Adlershof, beamline 14.2, λ = 0.91841 Å) used for data deposition of ALR2-**52** an additional

dataset was collected with our inhouse source (Incoatec IμS microfocus, λ = 1.5418 Å). Also in this dataset the carboxylic acid group was already missing in the resulting $F_o$-$F_c$ difference electron densitiy at a σ level of 3.0. Consequently, radiation damage can be excluded. The exchange of the nitro function by a carboxylic acid group leads to a surprising change in the binding mode. **51** addresses the specificity pocket and forms an H-bond with its nitro group to the Leu300 backbone. On the contrary, **52** directs the benzoic acid side chain towards the surrounding solvent area and fills the hydrophobic cleft adjacent to the unopened specificity pocket (Figure 5-4f, Figure 5-7f). Accordingly, the specificity pocket remains in its closed conformation even though the carboxylate group should be better suited to establish the H-bonds in the specificity pocket compared to the nitro group which is often classified as a weak H-bond acceptor. To enable this changed binding mode the bond between the central phenyl ring and the peptide bond is flipped over compared to **51**. In consequence, the intraligand H-bond between the amide nitrogen and one of the oxygen atoms of the acetic acid carboxylate group of **52** is ruptured. This internal H-bond was discussed in literature as crucial for high affinity binding.[43] The acetic acid head group of **52** adopts two different conformations in the anion binding pocket. Conformation A occurs with the minor occupancy (28%). It aligns well with the conformation found for the carboxylate group in all the other members of this series, accordingly a similar H-bonding pattern is established. Conformation B is the predominant one (49%). Here the carboxylate head group is shifted and forms an H-bond with one of its oxygen atoms to Tyr48OH (2.6 Å) whereas the other one contacts to His110NE2 (2.7 Å). At this position the second oxygen atom is too remote to make a direct H-bond to Trp111. Instead water W17 is captured (2.8 Å) and mediates such a contact. The existence of two conformations of the carboxylate group of the inhibitor and the entrapment of a water molecule at similar position was described previously for the binding mode of **52a** (Figure 5-2) at pH 8.0.[121] Owing to a sterical repulsion with W17 caused by one of the oxygen atoms of the carboxylate group in conformation A, the ligand can only entrap the water in conformation B. Its occupancy was therefore set to the same value as that of conformation B of the carboxylate group. W17 forms in turn an H-bond to Trp111NE1 (2.7 Å). Furthermore, W17 is in H-bond distance to the carbonyl oxygen atom of the central amide bond of **52** (2.8 Å). Supposedly, the generated interaction network by one of the carboxylate oxygens of **52** in conformation B, entrapment of W17 and the involvement of the peptidic oxygen of **52**

can, at least in parts, compensate for the loss of the internal H-bond between the carboxylate and the peptidic group of **52**. W2 is visible with conformations A and B. In both cases it makes H-bonds to the oxygen atom of the central amide bond (A: 3.3 Å, B: 2.5 Å). The amide nitrogen of **52** forms an H-bond to W16 (2.9 Å). W16 is found in an adjacent position in all the other described complexes, however, without establishing an H-bond to the ligand (not shown in Figure 5-7a-e). Thus, compared to the other described ligands, no additional water is trapped to establish this network. The terminal phenyl ring of the benzoic acid moiety of **52** undergoes an aromatic edge to face contact with the phenyl ring of Phe122 (~4.6 Å). Additionally, it forms a van der Waals contact to the aromatic moiety of Trp219 (3.8 Å). Compared to ALR2-**47** the residues 299-309 are slightly repositioned as similarly observed in ALR2-**48**. Here the RMSD, calculated with ProFit,[152, 153] for the residues 299-309 is 0.56 Å which is 4.4 times larger than the RMSD of the remaining protein residues. As shown in Figure 5-5a, this loop adopts a similar position as in ALR2-**48**.

### 5.4.10 The water networks around the surface of the specificity pocket

For the analysis of the water networks around the binding pocket of ALR2 the water molecules in contact with the adjacent symmetry mates in the packing were included. All water molecules with a distance not larger than 10 Å to **50** which was used as a reference ligand were included. Only those water molecules which are influenced by the binding of the different ligands are shown for the sake of clarity in Figure 5-7. For the same reason we focus on the first solvation shell. The oxygen atoms of the central peptide bond of the ligands **47**-**51** and **53** form an H-bond to water W1 (2.3-2.9 Å, Figure 5-4a-f). The latter is connected to a water cluster which consists of two connected five-membered ring systems next to the surface of Phe122, including W4, W5, W6, W7, W8, W9, W10 and W11. This cluster is fully visible in ALR2-**48** and ALR2-**51** (Figure 5-7b and e). The water molecules are stabilized by H-bonding contacts to residues of the protein (Phe122, Ser302). All here described complex structures crystallize in the spacegroup $P2_1$. Due to the spatial packing in the crystal a close symmetry mate of ALR2 comes close to the active site of the protein. The described water network in the crystal structures is stabilized by the OH and the amide group of Ser282 contributed by the next symmetry mate. Since the changes in the water networks which are induced by the different ligands are found more distal from the

latter residue the stabilization by the next symmetry mate should not take significant influence on the discussed water structures. In ALR2-**47** (Figure 5-7a) and ALR2-**48** (Figure 5-7b) W2 and W3 are mediating an H-bonding network to the amide nitrogen of Ala299 and the ligand. Remarkably in ALR2-**47** the NH of the peptide bond is used for the interaction to W2, whereas in ALR2-**48** the carbonyl oxygen of the peptide bond is used for the same interaction. In consequence in ALR2-**47** the water chain is deeper buried in the cleft between Trp219 and Leu300, in ALR2-**48** it is pushed upwards because the propynyl side chain of **48** fills the cleft. In the final refinement cycle of the complex structure of **47** we decided to not assign a water molecule at the position corresponding to that of W11 of ALR2-**48** and ALR2-**51**. Nevertheless, difference electron density suggests that W11 is at least partially occupied in ALR2-**47**. The $F_o$-$F_c$ difference electron density for W14 is shown in Figure 5-7c at a sigma level of 3.0. Reasonable difference electron density was visible but placing a water molecule led to sterical repulsions with the side chain carbon of Ser302. Summarized the binding modes of **47** and **48** induce a water pattern which involves a similar number of water molecules next to the surface of the specificity pocket. However, significant changes in the water structures are observed.

The terminal phenyl ring of **52** interferes with the water network next to Phe122 and Leu300. Nevertheless, the water molecules present in ALR2-**47** and ALR2-**48** can also be found in ALR2-**52**, however at shifted positions to avoid steric repulsions with the terminal phenyl ring of **52** (Figure 5-7f). W13, which is visible in ALR2-**50** and ALR2-**51**, but not in ALR2-**47** and ALR2-**48**, can also be assigned in ALR2-**52**. It forms H-bonds to Ser302OH (3.1 Å) and the amide nitrogen of the same residue (3.0 Å). Furthermore, W13 establishes an H-bond with W3 (3.0 Å) and van der Waals contacts (3.5 Å) to the terminal phenyl ring of **52**. W13 and W17, as already described above, are water molecules not detectable in the water networks around **47** and **48**, but in complex with **52**. All three ligands do not address the specificity pocket.

In ALR2-**51** (Figure 5-7e) W1, W12, W2 and W3 mediate a contact between the oxygen atom of the central peptide bond of **51** and the amide nitrogen of Ala299. W12 is additionally recruited to the network to bridge across the hydrophobic cleft formed by the side chain of Leu300. W13 is stabilized by H-bonds to the amide NH of Leu301 and to W11. It cannot be detected in ALR2-**47** and **48**.

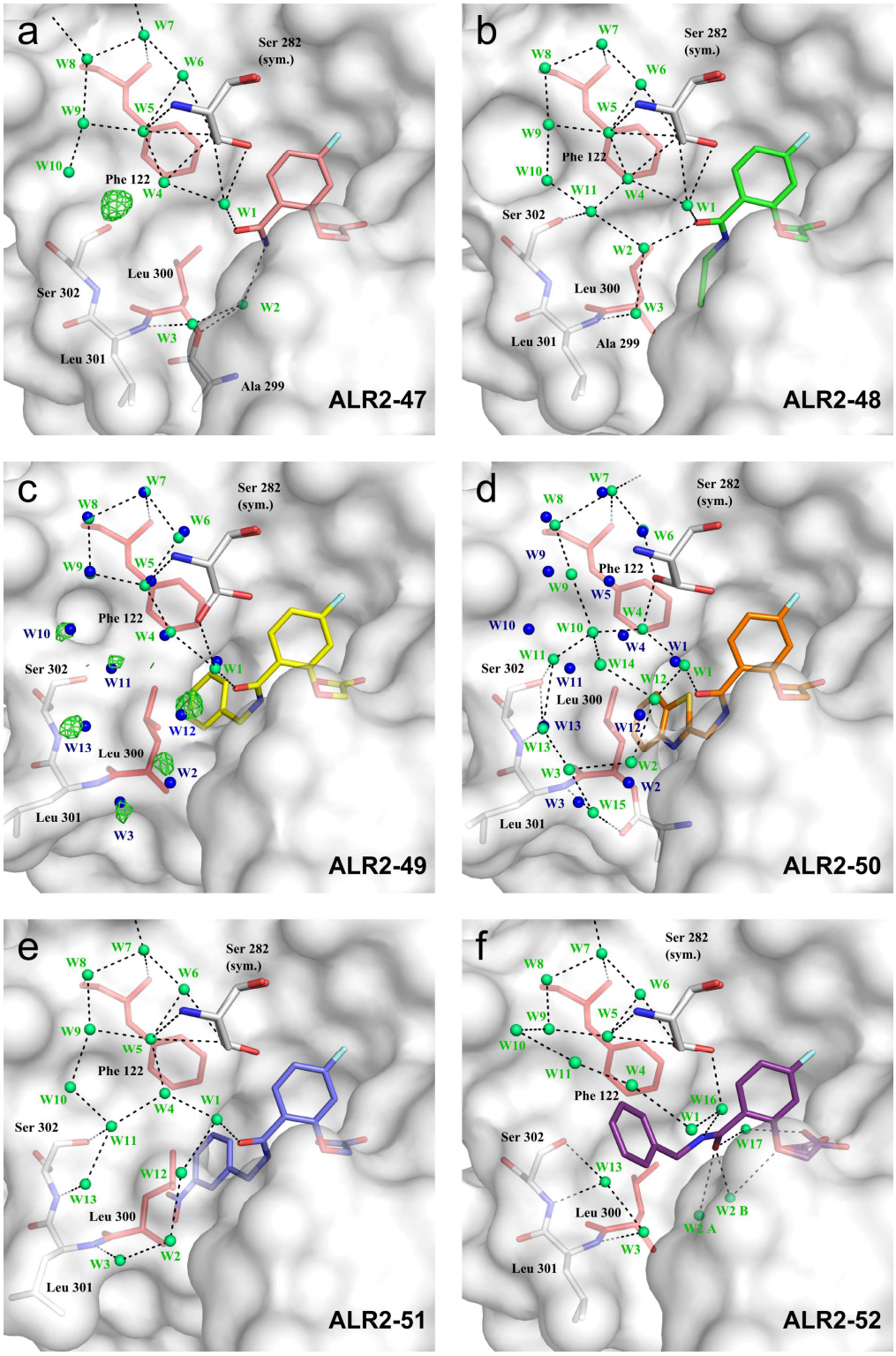



**Figure 5-7** Detailed figure caption on next page.

**Figure 5-7 Representation of the binding mode of 47-52.** The protein surface is shown in gray. Nitrogen atoms are colored in blue, sulfur atoms in yellow and oxygen atoms in red. Carbon atoms of the protein residues are shown in gray. H-bonds distances between 2.3 and 3.3 Å are drawn as black-dotted lines. (a) ALR2-**47**: Carbon atoms of **47** are shown in light red. The $F_o$-$F_c$ density at the position which corresponds to W11 is shown as a green mesh at a sigma level of 3.0 sigma. (b) ALR2-**48**: Carbon atoms of **48** are shown in green. (c) ALR2-**49**: Carbon atoms of **49** are shown in yellow. Water molecules which were placed in this structure are shown as green spheres. The water molecules from ALR2-**51** are shown as blue spheres. The $F_o$-$F_c$ difference electron densities from the ALR2-**49** dataset which correspond to the positions of water molecules W2, W3 and W10-W13 of ALR2-**51** is shown as a green mesh at a level of 3.0 σ. (d) ALR2-**50**: Carbon atoms of **50** are shown in orange. The water molecules from ALR2-**51** are shown as blue spheres. (e) ALR2-**51**: Carbon atoms of **51** are shown in blue. (f) ALR2-**52**: Carbon atoms of **52** are shown in purple.

In summary, the structures showing the specificity pocket in closed conformation, entrap two additional water molecules (W12, W13) which become involved in the H-bond network next to the specificity pocket compared to the complexes in open state. The partial occupancy of the ligand in ALR2-**49** along with the superimposed conformer of the *holo*-enzyme reduce the accuracy of the electron density and make it difficult to fully place the water molecules which would correspond to W2, W3, W10, W11, W12 and W13 in ALR2-**51**. Figure 5-7c shows apart of reliably assigned water molecules in ALR2-**49** (green spheres) the residual $F_o$-$F_c$ difference electron density in this complex together with the positions of the water molecules as formed in the ALR2-**51** complex (blue spheres) at a 3.0 σ level. The image shows that water molecules of ALR2-**51** coincide well with the residual difference electron density peaks in ALR2-**49**. Consequently, the water network observed for ALR2-**49** can be assumed to be identical to that of ALR2-**51**.

As discussed above, **50** induces an open conformer deviating from that found with **49** and **51**. As a consequence also the water network changes slightly (Figure 5-7d). The corresponding water molecules from ALR2-**51** are shown as blue spheres. Due to the different orientations of the side chain of Leu300 in ALR2-**50** the water molecules W4 and W11 are shifted compared to ALR2-**51**. A direct H-bond between W4 and W11 cannot be formed any longer . As a consequence W10 moves its binding position and forms in ALR2-**50** H-bonds to W4 (2.4 Å) and W11 (2.3 Å). The pivotal water W5 in ALR2-**51** is no longer stabilized. Instead W14, not observed in the other complexes, is recruited in ALR2-**50**. It seals a five-membered water cluster consisting of W1, W4, W10, W12 and W14 (Figure 5-7d). Furthermore, in this complex the backbone conformation next to Leu300 allows stabilization of W15 via H-bonds to the carbonyl oxygen of Ala 299 and the amide NH of Leu301 (Figure 5-7d). Overall, despite significant changes in the

water network, only one additional water molecule is involved in the H-bonding network to fix **50** compared to **51**.

### 5.4.11 Discussion of the binding mode and the surrounding water molecules of 53

Compared to **51**, **53** has an additional 4-phenyl substituent attached to the 2-(2-carbamoyl-phenoxy)acetic acid scaffold and the 5-F-substituent is missing (Figure 5-2). Figure 5-8a and b show the binding mode of **53**. The alignment of **53** and **51** shows very similar binding geometry for both ligands (Figure 5-8c). Consequently the main interactions with the protein residues are similar to the those found for **51**. Two main differences are detected. First, the side chain of Phe122 shifts in ALR2-**53** compared to ALR2-**51** to enable an edge-to-face stacking (3.8 Å) with the additional 4-phenyl ring of **53**. Secondly, this additional phenyl ring of **53** interferes with the water network which is detected in ALR2-**51**. For comparison purposes, the binding mode of **51** is shown in Figure 5-8d along with some selected water molecules which are represented as green spheres. In ALR2-**51** a contiguously connected chain of the water molecules (W1, W16, W18, W19, W20, W21 and W6) could be detected. W20 and W21 in the ALR2-**51** complex have been displaced in ALR2-**53** owing to the placement of the bulky 4-phenyl ring.

### 5.4.12 Differences in mean *B*-factors of the residues 299-309 of the C-terminal loop

The temperature factor (*B*-factor) reflects to some degree the residual mobility of an atom in a crystal structure.[154] Since the specificity pocket of ALR2 exhibits remarkable flexibility the temperature factors of the residues 299-309 were analyzed. These residues are part of the flexible loop region which borders the specificity pocket. The residues 310-315 were not included since they represent the C-terminus, which is known to possess a high flexibility independent of the conformational state of the pocket. In case of ALR2-**50**, ALR2-**51** and ALR2-**53** H-bonds between the ligand and the latter residues are formed which may take influence on their mobility. To consider the difference in the overall mobility of the involved residues across the series of structures the *B*-factors of the residues 299-309 were normalized with respect to the overall *B*-factor of the entire proteins and the *B*-factor ratio was mutually compared (Table 10-3 in the Appendix, numbers in brackets). This correlation was previously applied to compare residual mobilities.[155, 156] The *B*-factor ratios of the residues 299-309 show remarkable differences

between the different protein ligand complexes (1.2- 2.0) whereas of the ligands this property exhibits rather low and conserved values (0.8- 1.0).

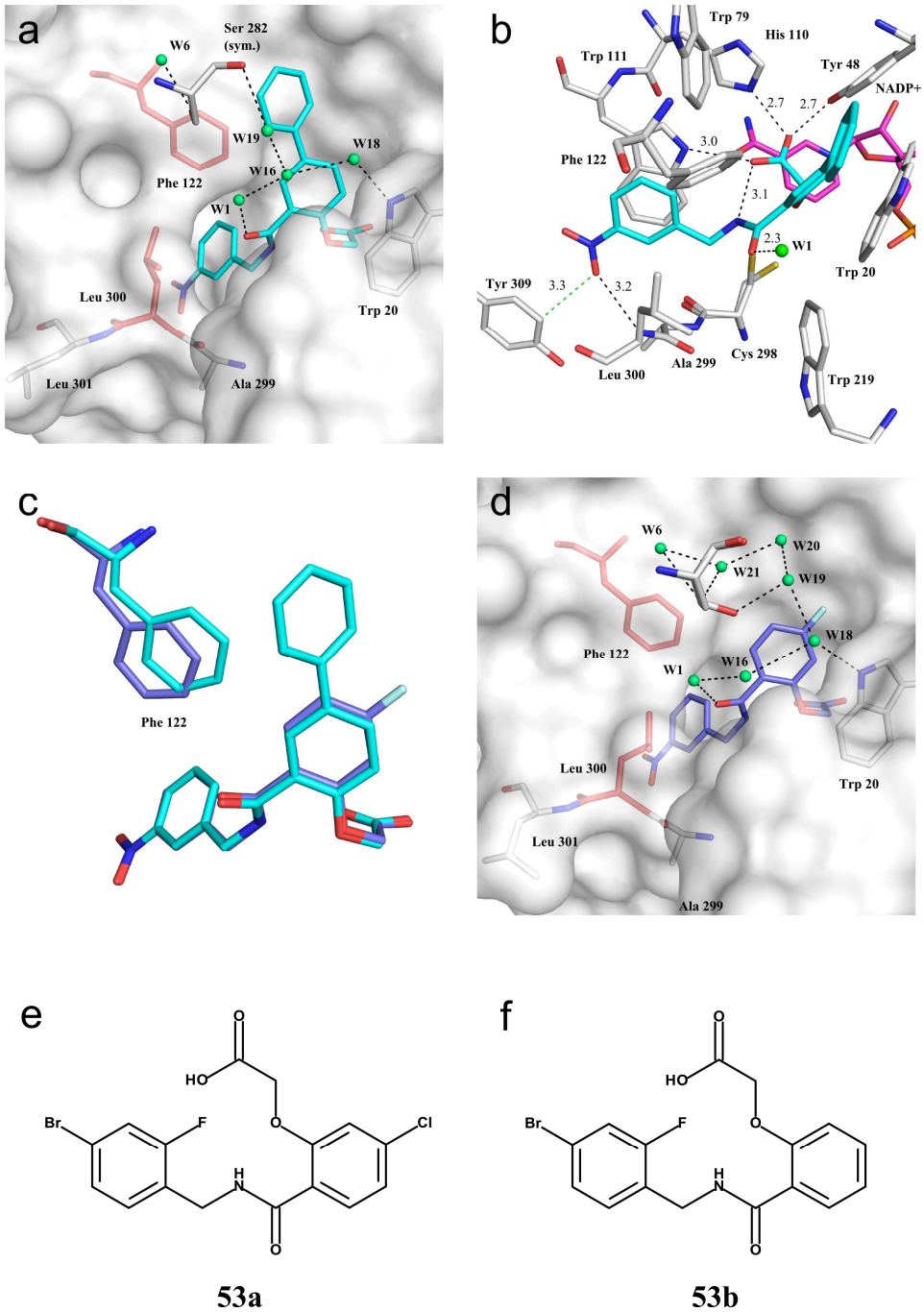



**Figure 5-8** Detailed figure caption on next page.

**Figure 5-8** (a) Binding mode of **53**. The surface of the protein is shown in gray. All atoms of Phe122 and Leu300 are shown in red. The carbon atoms of the remaining protein residues are shown in gray and of **53** in cyan. Oxygen atoms are shown in red, nitrogen atoms in blue, sulfur atoms in yellow and fluorine atoms in light blue. Water molecules are represented as green spheres and H-bonds as black-dotted lines. Only water molecules which are discussed in the text are shown. This color code is maintained during the following presentations. (b) Binding mode of **53** as a stick representation. The carbon atoms of $NADP^+$ are shown in magenta. The non-classical H-bond between one of the oxygen atoms of the nitro function of **53** and one of the hydrogen atoms of the phenyl ring of Tyr309 is drawn as a green-dotted line. (c) Mutual alignment of ALR2-**51** and ALR2-**53**. For ALR2-**51** the carbon atoms of Phe122 are shown in blue, for ALR2-**53** in cyan. **51** and **53** are aligning perfectly. Phe122 shifts in ALR2-**53** compared to ALR2-**51**, presumably to allow for an edge-to-face stacking to the attached 4-phenyl ring of **53**. (d) Binding mode of **51**. Carbon atoms of **51** are colored in blue. The water molecules W20 and W21 of ALR2-**51** are displaced by the additional phenyl substituent of **53**. (e) and (f) Chemical structures of **53a** and **53b**.

### 5.4.13 ITC measurements

The results of the ITC measurements are listed in Table 5-1 and summarized in Figure 5-9. The measurements were performed in HEPES buffer. An overlaid protonation effect can be assumed as constant, as explained in Section 5.7.2. The enthalpic data is not corrected for the superimposed protonation effect but cancels out in a relative comparison of the enthalpy values. The thermodynamic profiles for **47**-**50** result from direct ITC measurements in HEPES buffer. To obtain data for **51**, **50** and **53** the performance of displacement titrations was necessary. The experimental setup and the validation of the results of the displacement titrations are documented in Section 5.7.2.

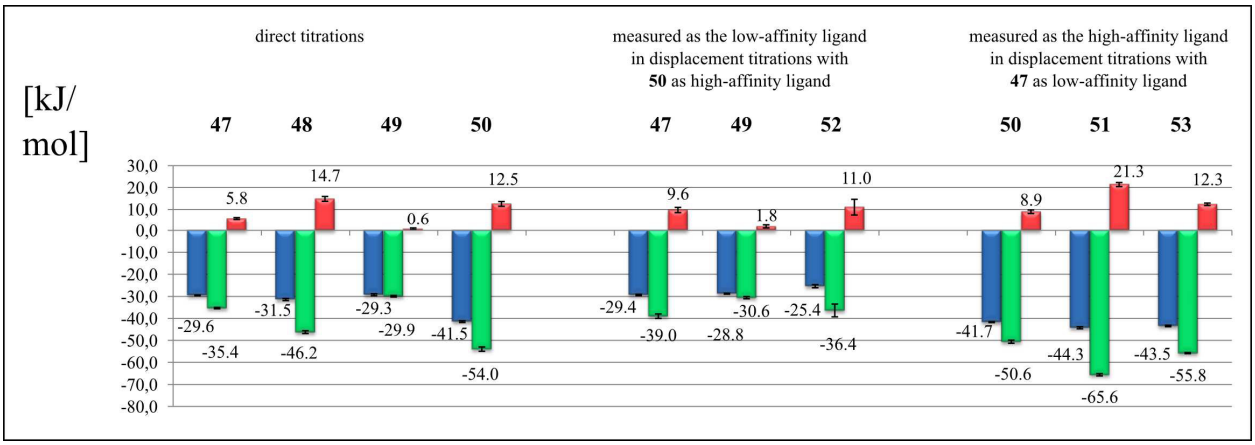



**Figure 5-9 Thermodynamic data.** The Gibbs free energy ($\Delta G^0$) is shown in blue, the enthalpic contribution ($\Delta H^0$) in green and the entropic contribution ($-T\Delta S^0$) in red. The measurements were performed in HEPES buffer and are not corrected for the overlaid protonation effect.

**Table 5-1 Thermodynamic data of the measurements performed in HEPES buffer.**

| | $K_D$ [µM] | $\Delta G^0$ [kJ/mol] | $\Delta H^0$ [kJ/mol] | $-T*\Delta S^0$ [kJ/mol] |
|---|---|---|---|---|
| | | **Direct Titration** | | |
| **47** | 6.6 ± 0.4 | -29.6 ± 0.2 | -35.4 ± 0.3 | 5.8 ± 0.4 |
| **48** | 2.8 ± 0.1 | -31.5 ± 0.5 | -46.2 ± 0.7 | 14.7 ± 1.1 |
| **49** | 7.5 ± 1.4 | -29.3 ± 0.4 | -29.9 ± 0.4 | 0.6 ± 0.4 |
| **50** | 0.054 ± 0.008 | -41.5 ± 0.4 | -54.0 ± 1.0 | 12.5 ± 1.1 |
| | | **Displacement Titration with 50 as high affinity ligand** | | |
| **47** | 7.2 ± 0.9 | -29.4 ± 0.3 | -39.0 ± 1.0 | 9.6 ± 1.2 |
| **49** | 9.1 ± 0.7 | -28.8 ± 0.2 | -30.6 ± 0.5 | 1.8 ± 0.7 |
| **52** | 36.2 ± 11.4 | -25.4 ± 0.7 | -36.4 ± 2.9 | 11.0 ± 3.6 |
| | | **Displacement Titration with 47 as low affinity ligand** | | |
| **50** | 0.050 ± 0.005 | -41.7 ± 0.2 | -50.6 ± 0.7 | 8.9 ± 0.8 |
| **51** | 0.018 ± 0.003 | -44.3 ± 0.5 | -65.6 ± 0.5 | 21.3 ± 0.9 |
| **53** | 0.024 ± 0.003 | -43.5 ± 0.3 | -55.8 ± 0.2 | 12.3 ± 0.6 |

## 5.5 Discussion

### 5.5.1 ITC experiments

#### 5.5.1.1 Comparison of structural and thermodynamic data of 47 and 48

The crystal structure analyses reveal very similar binding modes for **47** and **48**. In both structures the specificity pocket is found in the closed state. Even though differences in the surface water layer are found, the same amount of water molecules is involved in the interaction pattern between the ligands and the residues adjacent to the specificity pocket. The main difference between both ligands is the accommodation of the propynyl side chain of **48** in the rather hydrophobic cleft formed between Leu300 and Trp219. Furthermore, the latter ligand pushes the Ala299-Tyr309 loop out of way so that its hydrophobic propynyl substituent can be accommodated. **48** shows a slightly more favorable Gibbs free energy ($\Delta\Delta G^0$= -1.9 kJ/mol) than **47**. The $\Delta\Delta G^0$ difference is the result of a significantly higher exothermic binding of **48** ($\Delta\Delta H^0$= -10.8 kJ/mol) which is not fully compensated for by an entropically less favorable binding of **48** ($-T\Delta\Delta S^0$= +8.9 kJ/mol) compared to **47**. The *B*-factor ratio of the Ala299-Tyr309

stretch is much lower for **48** than for **47** (1.3 compared to 2.0). Presumably this is the consequence of the displacement of residues 299-309 which seem to be relocated to a less flexible state, leading to an unfavorable increase in the entropic binding portion. ALR2-**47** was obtained by soaking, ALR2-**48** by cocrystallization. Unlikely the different crystallization procedures provide an explanation for the deviating *B*-factor ratios of Ala299-Tyr309 stretch of these two complex crystal structures. For example, ALR2-**51** was, similarly to ALR2-**47**, obtained by soaking but in contrast to ALR2-**47** it exhibits a very low *B*-factor ratio for the latter residues. The binding of the propynyl side chain in the cleft is accompanied by a strong favorable increase of -10.8 kJ/mol of the enthalpy. On first glance this is surprisingly since the molecular portion only forms hydrophobic interactions. In this comparison it should not be neglected that also significant changes in the water structures are observed.

#### 5.5.1.2 Comparison of structural and thermodynamic data of 49 and 47

**49** exhibits with its unsubstituted phenyl ring the simplest moiety of all ALR2 inhibitors that succeeds in opening the specificity pocket. As discussed, the terminal phenyl ring of **49** does not form any other interactions in the specificity pocket apart from the π-π-stacking. In consequence with respect to **47**, which does not open the specificity pocket, **49** should allow to investigate the changes of the thermodynamic profile directly linked with the opening of the pocket. As discussed above, the $F_o$-$F_c$ difference electron density at 3.0 σ suggests that **49** entraps two additional water molecules (W12, W13) in surface network compared to **47**. The nitrogen of the terminal amide group of the latter ligand allows an H-bond to be formed to W2, a contact impossible to realize with **49**. Since W2 participates in a water-mediated H-bond to the amide group of the Ala299/Leu300-backbone, **47** establishes an additional contact to the protein residues, which **49** most likely compensates via W12. The predominant interaction formed by **49** in the specificity pocket is the π-π-stacking to Trp111. The affinity data show that both compounds bind with similar $K_D$ values (**47**: 6.6 µM, **49**: 7.5 µM) and Gibbs free energies ($\Delta\Delta G^0$= +0.3 kJ/mol). We therefore conclude that the opening of the specificity pocket and stacking interaction inside this pocket is not reflected by an increase in potency. **47** displays a more favorable enthalpic portion ($\Delta\Delta H^0$= -5.5 kJ/mol) nearly compensated by a less favorable entropic portion ($-T\Delta\Delta S^0$= +5.2) than **49**. The *B*-factor ratios of Ala299-Tyr309 of both complexes (1.9/2.0) are in the same range, indicating that these residues possess in both crystal

structures similar residual mobility. Thus, the sole opening of the specificity pocket does not restrain the flexibility of the 299-309 loop. First, the two different water networks should be contrasted with respect to the thermodynamic profiles. Even though the contact to W2 differs on first glance for **47**, in ALR2-**49** the corresponding W2 is stabilized via the water chain W1-W12-W2-W3. As a result for both complexes the contacts reach out to the backbone amide group of Leu300/Leu301. We therefore believe that W2 has similar influence on the thermodynamic profiles of both ligands and hardly discriminates the deviating thermodynamic profiles. The entrapment of W12 and W13 in ALR2-**49** should lead in contrary to ALR2-**47** to an increase in the enthalpic signal and a loss in the entropic portion, however, the opposite is observed. Thus, either the effect of the water fixation is very weak or additional effects compensate the water contribution.

The specificity pocket of ALR2 is presumably in its closed state before being addressed and opened by the attached phenyl substituent of an inhibitor. This assumption is supported by the crystal structure ALR2-**49**. The residual $F_o$-$F_c$ difference electron density in this complex shows the ligand 80% populated but superimposed density for the closed conformation of the specificity pocket is indicated. We therefore suppose that this subpocket opens upon binding of the phenyl side chain of **49** during the soaking. If the specificity pocket is closed before a ligand binds, no solvent molecules will be bound inside this pocket. Accordingly, the classical hydrophobic effect,[146-148] which involves the displacement of ordered water molecules from a hydrophobic environment to the bulk water phase cannot be the explanation for the more favorable entropic portion of **49**. In contrast, the phenyl ring of the inhibitor forming most likely unfavorable contacts to water molecules in the bulk phase is favorably buried in the rather hydrophobic specificity pocket. Since the enthalpic signal of **49** is smaller than that of **47** the desolvation of the aromatic ring and the formed π-π-stacking with Trp111 takes overall an unfavorable influence on the binding enthalpy.

As the unsubstituted **47** and the analog phenyl derivative **49** exhibit nearly identical affinity, the opening of the pocket along with the accommodation of the phenyl ring is linked to a tiny change in Gibbs free energy. A partitioning in compensating enthalpically unfavorable and entropically beneficial contributions is found overall. Further splitting into individual contributions is difficult

as **49** (i) entraps two additional water molecules (usually an enthalpically favorable but entropically unfavorable step), (ii) has to drive the protein into a different protein conformation (possibly linked to a small energy barrier as the closed form is more stable), (iii) will require a contribution to desolvate the aromatic portion (a step which might be enthalpically unfavorable but entropically beneficial) and (iv) weak but energetically favorable π-π stackings are experienced. Clearly arguments can be found to relate these steps to a more enthalpic or entropic signature, however it appears impossible to quantify them relative to each other to reveal the profile for the entire step of the pocket opening.

#### 5.5.1.3 Comparison of structural and thermodynamic data of 50 and 49

The increased potency of **50** compared to **49** ($\Delta\Delta G^0$= -12.2 kJ/mol) is the result of the much more favorable enthalpic contribution ($\Delta\Delta H^0$= -24.1 kJ/mol) which is only partly compensated by a higher entropic price ($-T\Delta\Delta S^0$= +11.9 kJ/mol). The favorable increase in the enthalpic portion can be related to the formed H-bond between the nitrogen of the benzothiazolyl ring of **50** and the backbone of Leu300, even though this gain appears surprisingly large. This formed H-bond restraints the flexibility of the residues 299-309, leading to a significant smaller *B*-factor ratio for these residues (1.4 compared to 2.0) which supposedly explains the more unfavorable increase in the entropic portion of **50**.

#### 5.5.1.4 Comparison of structural and thermodynamic data of 51 and 49

The comparative analysis of **51** and **49** allows to dissect the influence on the thermodynamic parameters of the *meta* nitro group of the phenyl ring which addresses the specificity pocket of ALR2. ALR2 inhibitors bearing this group have been described previously as strong enthalpic binders[8] but the portion which can be contributed to the attached *meta* nitro group could not be quantified so far. The thermodynamic profile of **51** could only be measured in a displacement titration. In a recent study[157] we could show that displacement titrations can involve a more complex behavior making the assumption of a full additivity in all steps of the displacement process questionable. In the current case we could titrate **47** and **49** directly and by displacement with a strong binding ligand, whereas **50** could be studied directly and by displacing a weak ligand used to preincubate the protein. As reported in Section 5.7.2, the Gibbs free energy is well reproduced (≤0.5 kJ/mol), the partitioning in enthalpy and entropy is affected by larger

modulations but below 4 kJ/mol. Compared to **49**, **51** forms an additional H-bond via its nitro function to the backbone of Leu300 and a non-classical H-bond to the CH of the phenyl ring of Tyr309. The formation of these interactions are accompanied by a surprisingly large gain in affinity ($\Delta\Delta G^0$= -15.0 kJ/mol), overwhelmingly due to an enthalpic gain ($\Delta\Delta H^0$= -35.7 kJ/mol). The deviating *B*-factor ratios for Ala299-Tyr309 (ALR2-**49**: 2.0, ALR2-**51**: 1.2) indicate that the formed H-bond strongly decreases the flexibility of the residues 299-309. This explains the loss in the entropy of **51** compared to **49**.

#### 5.5.1.5 Comparison of structural and thermodynamic data of 50 and 51

**50** and **51** are both forming H-bonds in the specificity pocket to the backbone of Leu300 with similar distances (3.1 Å/ 3.2 Å). The thermodynamic profiles of both ligands slightly differ in affinity ($\Delta\Delta G^0$= -2.6 kJ/mol) and factorize quite strongly in $\Delta\Delta H^0$= -15.0 kJ/mol and $-T\Delta\Delta S^0$= +12.4 kJ/mol. Obviously, the formed H-bond between the oxygen of the nitro group of **51** and the Ala299-Leu300 backbone NH is accompanied by a stronger increase in the enthalpic signal compared to the H-bond between the basic nitrogen of the benzothiazolyl moiety of **50**. Furthermore, the nitro group of **51** forms a non-classical H-bond to a hydrogen of the phenyl ring of Tyr309 which may further increase the exothermic signal. Presumably the enthalpic desolvation costs for the nitrophenyl group, heavily involved in contacts to the protein, are less than for the benzothiazolyl moiety. The *B*-factor ratio of Ala299-Tyr309 (**51**: 1.2, **50**: 1.4) indicates that **51** locks the residues 299-309 more firmly in their position as **50** does. This is also supported by the entropic portion which is less favorable for **51** than for **50**. Additionally, in ALR2-**50** the backbone next to Ala299 adopts a slightly different conformation as in ALR2-**51**. This conformation might be energetically less favorable which would additionally decrease the enthalpic signal for ALR2-**50**. The stabilization of water molecule W15 in ALR2-**50**, which should be accompanied by a favorable increase in the enthalpic signal and a loss in the entropic part, is obviously overcompensated by the different interactions of both ligands in the specificity pocket. Furthermore, the deviating arrangements of the water networks next to the specificity pocket which were discussed above could lead to deviating thermodynamic signatures.

### 5.5.1.6 Comparison of structural and thermodynamic data of 52 and 47

The less potent **52** and **47** show different thermodynamic profiles ($\Delta\Delta G^0$= +4.0 kJ/mol with $\Delta\Delta H^0$= +2.6 kJ/mol and $-T\Delta\Delta S^0$= +1.4 kJ/mol). The decrease in the enthalpic term for **52** results, at least partially, in the loss of the internal H-bond between the amide nitrogen of the ligand and one of the oxygen atoms of its acetic acid carboxylate group. Furthermore, the altered binding mode of the polar head group of **52** in the anion binding pocket should also be reflected in the thermodynamic signature. Two additional water molecules are involved in the water network stabilizing **52** compared to that of **47** (W13, W17). Furthermore, the terminal phenyl ring of the latter ligand contributes with favorable contacts to Phe122. The *B*-factor ratio for the 299-309 stretch (1.7/ 1.9) suggests that **52** decreases flexibility of this loop slightly stronger than **47**. As described above, **52** displaces the residues 299-309 similarly to **48**, thus reducing mobility, but less dominantly as ALR2-**48**. This could explain the unfavorable entropic contribution of **52**.

Surprisingly is the fact that **52** places its terminal benzoic acid moiety in direction of the surrounding solvent and not into the specificity pocket as found for the isosteric nitrophenyl group of **51**. A possible explanation could provide the much more costly desolvation of acidic benzoic acid group. At the adopted position it remains partly solvent exposed whereas accommodation in the specificity pocket would require complete desolvation.

### 5.5.1.7 Comparsion of the structural and thermodynamic data of 51 and 53

The $IC_{50}$ values for **51** and **53** were reported as similar (8 nM).[43] Our ITC study comes to comparable results since the $K_D$ values (**53**: 0.024 µM; **51**: 0.018 µM, $\Delta\Delta G^0$= -0.8 kJ/mol) show only a slightly decreased affinity for **53** compared to **51**. Interestingly, the thermodynamic profiles of both ligands, which we obtained in a displacement titration using **47** as a weak, competing ligand, differ significantly: $\Delta\Delta G^0$= -0.8 kJ/mol, $\Delta\Delta H^0$= -9.8 kJ/mol and $-T\Delta\Delta S^0$= +9.0 kJ/mol. The data published by Van Zandt et al.[43] allow to estimate the influence of a halogen substituent in 5-position at the 2-(2-carbamoyl -phenoxy)acetic acid scaffold. A ligand which carries a 5-chlorine substituent (**53a**, $IC_{50}$: 30 nM, Figure 5-8e) has an $IC_{50}$ value which is about six times smaller as the analogue without a halogen substituent at this position (**53b**, $IC_{50}$: 176 nM, Figure 5-8f). On the other hand, the exchange of a clorine atom at this position (**46**, $IC_{50}$: 6 nM) by a fluorine atom (**51**, $IC_{50}$: 8 nM) does not influence the affinity significantly.[43]

Obviously, in the here described case the additional phenyl ring in 4-position of **53** can largely compensate for the missing halogen atom in 5-position regarding the free energy of binding, even if the enthalpic and entropic partitioning differs for both ligands. As already mentioned the additional phenyl ring of **53** replaces the water molecules W20 and W21. These water molecules are released to the bulk solvent phase and thus their degrees of freedom will increase resulting in a more favorable entropic term compared to **51**. In contrast, the enthalpic portion is less favorable presumably as a result of the disturbed water structure leading to a loss of H-bond contacts in the system. This might lead to a more unfavorable enthalpic portion for **53** compared to **51** which cannot be compensated for by the additional edge-to-face stacking of the attached 4-phenyl ring of **53**.

## 5.6 Conclusions

Addressing the specificity pocket of ALR2 may provide the opportunity to design ALR2 inhibitors with pronounced selectivity over binding to ALR1. In this study we provide for the first time insights into the thermodynamic properties involved in the opening of this pocket. We determined high-resolution crystal structures of a narrow series of inhibitors which induce three different binding site conformers. The ITC data reveal large differences in Gibbs free energy, enthalpy and entropy. The comparison of **47** and **48** shows that ligands which possess similar binding modes can still show strongly deviating thermodynamic profiles. Thus, such profiles do not provide straightforward a conclusive answer whether the specificity pocket is opened or not. The comparison of **49** and **51** allowed to quantify the enthalpic and entropic contributions of the attachment of a *meta* nitro group. Facing the isosteric ligands **51** and **52**, only differing in a COOH by $NO_2$ exchange, the consequences of this displacement on the thermodynamic profile could be studied. The exchange leads surprisingly to an alternate binding mode of the ligand portion. **52** directs its terminal benzoic acid toward the surrounding solvent while the nitro-phenyl ring of **51** is deeply buried in the specificity pocket. This suggests that the desolvation penalty for a nitro group is significantly lower than for a carboxylic acid group. The comparison of **47** and **49** shows that the enthalpic signal is significantly reduced when part of the ligand addresses the specificity pocket without forming H-bonds inside the pocket. This suggests that some energy is needed to transform the protein residues into the open specificity pocket

conformer. The presented case study provides some insights into the thermodynamic properties involving the opening of a protein pocket. They will afford valuable input for other drug design projects with proteins also exhibiting pockets of transient nature.

## 5.7 Description of Experimental Procedures

### 5.7.1 Expression, purification, crystallization and structure determination

The expression and purification of ALR2 are described in Section 9.2.1, the crystallization in Section 9.2.3, the soaking in Section 9.2.3.7 and the structure determination in Section 9.2.4.

### 5.7.2 ITC experiments

The measurements were performed with an $ITC_{200}$ (GE Healthcare) at 298 K. For the direct titrations of **47**-**50** the cell was filled with a solution containing ALR2 (10 µM-190 µM), an excess of $NADP^+$ and 3% (v/v) DMSO in 10 mM HEPES buffer at pH 8.0. The syringe was filled with the same concentration of $NADP^+$ and DMSO and with the inhibitor (0.075-1.2 mM) in 10 mM HEPES at pH 8.0. 20-25 injections with a volume from 1.6–2.0 µL were done. An initial injection of 0.3-0.5 µL was excluded from the analysis.

It was shown previously that the enthalpic signals of carboxylate-type ALR2 inhibitors display a buffer dependency due to an uptake of approximately 0.8-0.9 mol protons per mol formed complex.[42] The latter data were measured with a MCS-ITC instrument of MicroCal in pyrophosphate, HEPES, Tricine and TRIS buffer. Recently we determined with the $ITC_{200}$ instrument a proton uptake of 0.6-0.7 per mole of formed complex for different types of ALR2 inhibitors bearing a carboxylate head group (unpublished results and Stefek et al.[127]) using HEPES, Tricine and TRIS buffer each at concentrations of 10 mM and at pH 8.0. The thermodynamic profile of two of the here investigated compounds, **48** and **50**, were measured exemplarily in the latter three buffer systems. The resulting enthalpies ($\Delta H^0_{obs}$) were plotted against the heat of ionization of the corresponding buffer system ($\Delta H^0_{ion}$). The values for $\Delta H^0_{ion}$ were taken from literature.[87, 88] A line of best fit was drawn and its interception with the y-axis indicated the enthalpy to be corrected for the protonation effect ($\Delta H^0_{bind}$). The slope denoted the molar quantity of protons that are entrapped or released according to the formation of one mole of complex. A positive slope, as observed here, indicates proton uptake. The resulting diagram

for the binding of **48** and **50** is shown in Figure 5-10, the values of $\Delta H^0_{ion}$ and $\Delta H^0_{obs}$ are shown in Table 5-2.

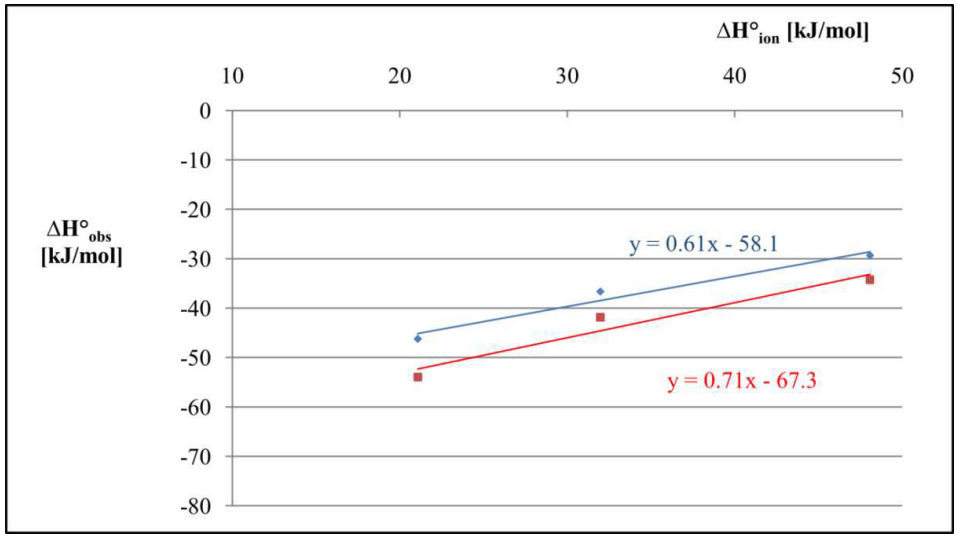



**Figure 5-10 Plot of the heat of ionization ($\Delta H^0_{ion}$) of the corresponding buffer system (HEPES, Tricine or TRIS) against the measured enthalpy ($\Delta H^0_{obs}$) of 48 (blue) and 50 (red).**

**Table 5-2 Values for the heat of ionization ($\Delta H^0_{ion}$) of the different buffer systems and the corresponding measured enthalpies ($\Delta H^0_{obs}$) of 48 and 50.**

| | | **48** | **50** |
|---|---|---|---|
| Buffer | $\Delta H_{ion}$ [kJ/mol] | $\Delta H_{obs}$ [kJ/mol] | $\Delta H_{obs}$ [kJ/mol] |
| HEPES | 21.07 | -46.2 | -54.0 |
| Tricine | 31.97 | -36.7 | -41.9 |
| TRIS | 48.07 | -29.4 | -34.3 |

The resulting proton uptake of approximately 0.6-0.7 mol per mol of formed complex is in agreement with our previous results. It was shown that the proton is picked up by Tyr48.[42] The moiety of the here investigated ligands which binds in proximity to this residue does not vary. The only substituent of the remaining ligands which could be involved in a proton transfer is the benzoic acid moiety of **52**. However, at pH 8.0 this group should be completely deprotonated. We assume a similar influence on the enthalpic signal should be given for all the compounds of the studied series.

Consequently, we refrained from correcting the enthalpy for the superimposed protonation effect since we are only interested in the relative differences between the investigated ligands. All measurements which led to the discussed thermodynamic profiles were performed in HEPES buffer.

The strong affinity of **51** and **53** and the weak affinity of **52** made the direct titration of these ligands difficult. Consequently, displacement titrations were chosen to obtain the thermodynamic data for the latter ligands. For the displacement titration of **51** and **53** the protein was saturated with **47** as a competing ligand with a weaker affinity. The same concentration of **47** was added to the syringe. The data for **47** ($K_{a,\ weak\ ligand}$, $\Delta H°_{weak\ ligand}$) were obtained by direct titrations of this ligand. Using equations 5-1 and 5-2, which were adapted from Sigurskjold,[158] $K_{a,\ strong\ ligand}$ and $\Delta H°_{strong\ ligand}$ for the high-affinity binder could be calculated.

$$K_{a,strong\ ligand} = K_{a,observed}\left(1 + K_{a,weak\ ligand}[weak\ ligand]\right) \quad (5\text{-}1)$$

$$\Delta H^{\circ}_{strong\ ligand} = \Delta H^{\circ}_{observed} + \left(\Delta H^{\circ}_{weak\ ligand}\ \frac{K_{a,weak\ ligand}[weak\ ligand]}{1+K_{a,weak\ ligand}\ [weak\ ligand]}\right) \quad (5\text{-}2)$$

The protein concentration varied between 23-42 µM. The required concentration of 90 µM **47** was determined in a preceding direct titration of this ligand. This concentration exceeds 2-4x the concentration of the protein and approximately 15x the $K_D$. To validate the experimental set up, **50**, which could also be measured in a direct titration previously, was titrated under the same condition. The $K_D$ value found for **50** by the direct titration (0.054 µM) and the displacement titration (0.050 µM) agreed very well. ΔH° (direct: -54.0 kJ/mol, displacement: -50.6 kJ/mol) and the $-T\Delta S^0$ (direct: +12.5 kJ/mol, displacement: +8.9 kJ/mol) differed only slightly. Furthermore, **51** was measured in a direct titration. The high c-values of the recorded ITC titration curves made the extraction of the $K_a$ value and thus the full thermodynamic profile difficult. However, ΔH° can be evaluated. The obtained value of -67.0 ± 2.5 kJ/mol agrees well with the value of -65.6 ± 0.6 kJ/mol which was obtained by the displacement titration.

**52** is a too weak binder to be measured in a direct titration. Here the inverse experimental design was chosen with **52** as a weak competing binder which was in turn replaced by **50**. Again, to evaluate the protocol, **47** and **49**, for which data from direct titrations were available, were titrated similarly. The applied protein concentration was 31 µM. Again, at first direct titrations of the weak ligands were performed to detect the concentration required to saturate the protein. This resulted in 90 µM (**47**, **49**) and 180 µM (**52**) which exceeds 3.9- 7.6 x the protein concentration and 5-15x the $K_D$. The same concentration of the weak binder was added to the syringe. The thermodynamic parameters of the weak binder, $K_{a,\ weak\ ligand}$ and $\Delta H°_{weak\ ligand}$, were calculated using the equations 5-3 and 5-4 which were adapted from Zhang et al.:[159]

$$K_{a,weak\ ligand} = \left(\frac{K_{a,strong\ ligand}}{K_{a,observed}} - 1\right) \frac{1}{[weak\ ligand]} \tag{5-3}$$

$$\Delta H^{\circ}_{\ weak\ ligand} = (\Delta H^{\circ}_{\ strong\ ligand} - \Delta H^{\circ}_{\ observed}) \left(1 + \frac{1}{K_{a,weak\ ligand}\,[weak\ ligand]}\right) \tag{5-4}$$

**49** led to very similar results for direct and displacement titrations ($K_D$ direct/ displacement: 7.5 µM/ 9.1 µM; ΔH° direct/ displacement: -29.9 kJ/mol /-30.6 kJ/mol; $-T\Delta S^0$ direct/ displacement: +0.6 kJ/mol/ +1.8 kJ/mol). The results for **47** differed slightly stronger ($K_D$ direct/ displacement: 6.6 µM/ 7.2 µM; ΔH° direct/ displacement: -35.4 kJ/mol /-39.0 kJ/mol; $-T\Delta S^0$ direct/ displacement: +5.8 kJ/mol/ +9.6 kJ/mol).

The presented comparisons of the thermodynamic data from the displacement titrations with the data obtained by direct titrations indicate that the displacement titrations are providing reliable results regarding the $K_D$ value while the error in the enthalpic portion seems to be slightly higher. However, a qualitative comparison of the enthalpic portions appears to be justified. The phenomena of differing results for ΔH° for direct and indirect titration has already been described elsewhere.[157] As a consequence in the discussion we are only comparing the results from direct or indirect titration among each other, if possible. For the comparison of **49** (direct titration) and **51**

(displacement titration) the thermodynamic data from different titration protocols were taken since the strongly deviating $K_D$ values did not allow for applying the same procedure.

The binding isotherms were integrated using NITPIC[160] and afterwards fitted with SEDPHAT.[161] Representative binding isotherms and fitted regression curves are shown in Figure 10-4 in the Appendix. All measurements were done at least in triplicate.

# 6 Kinetic and Structural Insights into the Mechanism of Binding of Sulfonamides to Human Carbonic Anhydrase

## 6.1 Introductory Remarks

The following chapter was prepared for submission to a scientific journal. In this chapter the main text and the supporting information are combined in contrast to the original draft. This project was a collaboration with Dres. Roberto Gaspari, Giovanni Bottegoni and Walter Rocchia from the group of Prof. Dr. Andrea Cavalli (Department of Drug Discovery and Development, Italian Institute of Technology, Genova, Italy), Dres. Daniel Schwarz and Jörg Bomke (Merck KGaA, Merck Serono Research, Small Molecule Platform/MIB, Darmstadt, Germany), Hans-Dieter Gerber and Prof. Dr. Andreas Heine (Philipps University Marburg). The author of this thesis performed the SPR experiments, determined the complex crystal structures of **54**-**58** with hCAII and contributed to the drafting of the manuscript.

## 6.2 Abstract

The binding of sulfonamides to human carbonic anhydrase II (hCAII) is a complex and long-debated example of protein-ligand recognition and interaction. In this study, we investigate the para-substituted n-alkyl and hydroxyethylene-benzenesulfonamides, providing a complete reconstruction of their binding pathway to hCAII by means of molecular simulations, surface plasmon resonance (SPR) measurements and X-ray crystallography experiments. The characterization of the binding pathway allows an unprecedented understanding of the structure-kinetic relationship in hCAII/benzenesulfonamides complexes, depicting a paradigmatic scenario for the multi-step binding process in protein-ligand systems.

## 6.3 Introduction

Human carbonic anhydrase II (hCAII) is an enzyme that catalyzes the transformation of carbon dioxide into bicarbonate. It is often used as a prototypical model system for biophysical studies[46] and medicinal chemistry applications.[47, 162] The catalytic site of hCAII consists of one $Zn^{2+}$ ion, which is coordinated by three histidine residues and a water molecule (or hydroxyl ion) in a distorted tetrahedral geometry. Sulfonamide-carrying ligands can inhibit hCAII by directly binding to the catalytic $Zn^{2+}$ ion of the protein and displacing the water molecule. While it is

generally believed that the hCAII-sulfonamide recognition process is multi-step, little is known about the structural features of the kinetically relevant intermediate configurations along the binding pathway. As in most processes of protein-ligand binding, this can be traced back to the transient and short-lived nature of the intermediate states, which cannot always be detected by experiments.

Our knowledge of hCAII-sulfonamide complexes relies upon extensive experimental characterization carried out over the last few decades, which has produced key results that must be taken into account when considering models of enzyme-inhibitor binding. Generally, the binding of sulfonamides to hCAII is a slow process, which considerably depends on the nature of the substituents present on the benzenesulfonamide core. This points to a multi-step mechanism, as a single-step association would be more compatible with a diffusion-limited process, characterized by fast kinetics and only marginally dependent on the chemical decorations of the binding compounds.[66]

The impact of the ligand substituents on the association kinetics was thoroughly investigated in the case of para-substituted n-alkyl benzenesulfonamides, where stopped-flow fluorescence (SFF) experiments[66, 67] showed that the observed on-rate, $k_{on}$, increases exponentially with the length of the alkyl chain. These results have been interpreted in terms of a minimal two-step process, although no direct evidence of an intermediate state was produced by SFF measurements. The dissociation constants of the alkyl-substituted sulfonamides in complex with the zinc-free hCAII displayed a very high correlation with the octanol-water partition coefficient, a fact which was used to infer the key role of hydrophobic interactions in stabilizing the putative intermediate state.[66] From a structural standpoint, compounds containing alkyl chains can establish favorable interactions with the hydrophobic patch at the enzymatic cavity, often referred to as the "hydrophobic wall".[163] The role of these has recently been confirmed by isothermal titration calorimetry (ITC) experiments,[58,164] molecular dynamics[164] and site-directed mutagenesis experiments.[165-167]

Further complexity to the binding mechanism is given by the charge state of the sulfonamide, which binds to the zinc ion in its anionic form, as reported by $^{15}$N-NMR[168] and recent neutron

diffraction studies.[169] Since benzenesulfonamides are usually weak acids, it is expected that these inhibitors initially bind to hCAII in their neutral form and only subsequently deprotonate, upon reaching the final $Zn^{2+}$-coordinated configuration. On the other hand, the water molecule bound to the $Zn^{2+}$ has a pKa value of 6.8[170] and is thought to be a hydroxyl ion in the active form of the enzyme.[46] Experimental data on the pH dependence of the association process have pointed to two possible mechanisms: either i) a neutral sulfonamide coordinates the $Zn^{2+}$ displacing the hydroxyl ion, or ii) a charged sulfonamide displaces the $Zn^{2+}$-bound water molecule.[171] No consensus has yet been reached on the atomistic mechanism of sulfonamide binding to hCAII. In this scenario, the aim of this work is to provide a complete reconstruction of the binding process of para-substituted n-alkyl and hydroxyethylene-benzenesulfonamides (**54**-**58**, see Figure 6-1) to hCAII. The dynamical properties of the system are investigated via large-scale molecular dynamics (MD) simulations, allowing the identification of relevant pre-binding conformations adopted by compounds **54**-**58** upon their entrance into the enzyme cleft. A comparison between the predicted stability of the pre-binding conformations and the results of surface plasmon resonance (SPR) measurements is also reported. MD simulations clearly point to a multi-step binding mechanism, as previously hypothesized by Taylor, King and Burgen[66, 67] and disclose the postulated pre-binding intermediate state. The role of the hydrophobic tail mounted on the sulfonamide scaffold clearly emerges, allowing a thorough characterization of the structure-kinetics relationships (SKR) of benzenesulfonamides. We then carry out metadynamics simulations to provide an in-depth characterization of the protein-ligand energetic landscape, which allows a description of the ligand-binding pathway up to its coordination to the $Zn^{2+}$-bound water, which is considered in our simulations as a hydroxyl ion. Subsequently, we estimate the environment-induced pKa shift of the compounds via a Poisson-Boltzmann equation-based model and then study the sulfonamide deprotonation by means of density functional theory calculations (DFT). These results are thoroughly discussed in light of the binding models recently reviewed by Krishnamurty et al.[46] To describe the entire binding pathway, we obtain high-resolution co-crystal structures for compounds **54**-**58** in complex with hCAII. They unravel atomistic details of the stable $Zn^{2+}$-bound sulfonamide/hCAII complexes and disclose the sulfonamide conformations as the final thermodynamic minima of the entire binding process.

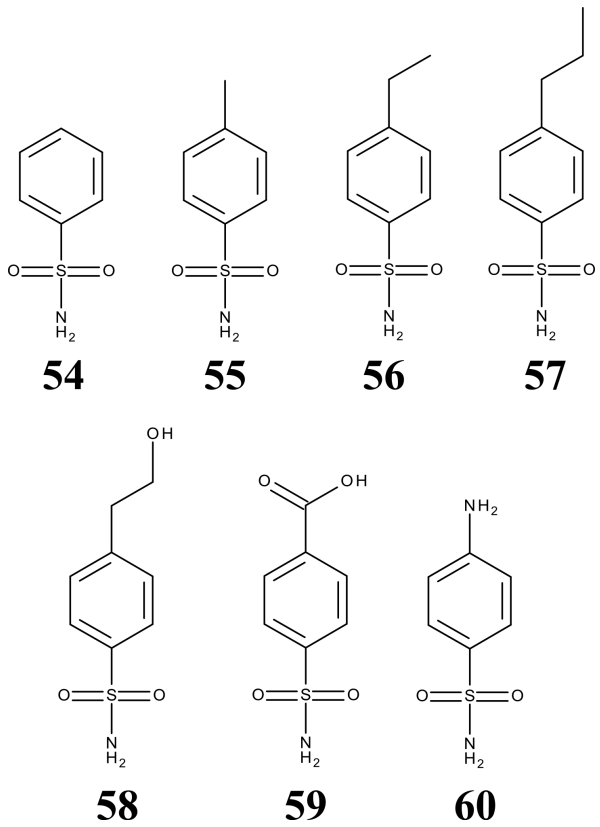



**Figure 6-1 Chemical structures of the para-substituted alkyl benzenesulfonamides (54-57), hydroxyethylene-benzenesulfonamide (58) and the reference compounds 59 and 60.**

**Table 6-1 The kinetic data obtained by SPR experiments.**

| | $K_D$ steady state[A] [nM] | $K_D$ kinetic[A] [nM] | $k_{on}$ [ x $10^5$ 1/Ms] | $k_{off}$ [1/s] |
|---|---|---|---|---|
| **54** | 1135.3 ± 196.2 | 1149.7 ± 101.7 | 1.4 ± 0.2 | 0.1630 ± 0.0164 |
| **55** | 331.7 ± 93.7 | 292.9 ± 46.3 | 2.0 ± 0.1 | 0.0582 ± 0.0073 |
| **56** | 125.4 ± 25.2 | 98.3 ± 7.4 | 4.7 ± 0.6 | 0.0461 ± 0.0033 |
| **57** | 60.6 ± 5.9 | 57.8 ± 5.4 | 7.8 ± 1.6 | 0.0454 ± 0.0129 |
| **58** | 428.6 ± 55.8 | 392.0 ± 51.8 | 1.6 ± 0.1 | 0.0643 ± 0.0079 |

[A] The definition of the steady-state and kinetic dissociation constants ($K_D$) are reported in Section 9.3.1.4.

## 6.4 Results

### 6.4.1 The role of hydrophobicity in the structure-kinetic relationship

It was previously reported, by means of SFF investigations,[66] that the association rates for **54**-**57** increase exponentially with the substituent length. Here, we report SPR measurements for the same ligand series as well as for the more hydrophilic one, compound **58**. We aim to test the effect of the methyl-hydroxyl substitution on the binding rates, to assess whether hydrophobicity or chain length is the major feature responsible for the observed $k_{on}$ trend.

### 6.4.2 Results of SPR measurements

The ligands **54**-**58** were measured one after another at four independent protein surfaces of 3700 RU, 3600 RU, 4900 RU and 4700 RU. The ligands **59** (4-CBS), **60** (sulfanilamide), **58** (94%) and **58** (98%) were measured consecutively at four independent protein surfaces of 4900 RU, 4800 RU, 4500 RU and 4500 RU. Mean values and the standard deviations for **54**-**58** are shown in Table 6-1**.** The corresponding results of the single measurements are listed in Table 6-2. The measured curves for each ligand are presented in Figure 10-5 - 10-8 in the Appendix. In preliminary tests, an impurity was discovered for compound **58**. It was purified to at least 94% in the first sample; a second purification step provided 98% purity. Two dilution series, one with the substance with 94% purity and a second based on 98% purity, were measured in four independent SPR experiments and the kinetic parameters for the binding to hCAII were determined. As shown in Table 6-3, the results are very similar for both purity levels. Thus, the reduced purity of **58** (94%) compared to the ligands **54**-**57** does not explain the deviating kinetic parameters. We validated our assay by measuring **59** and **60** as reference ligands and comparing our results with previously determined values by Myska[60] for the binding of these compounds to the bovine CAII (bCAII). Both compounds possess a very similar chemical structure compared to **54**-**58**. Two independent series of measurements with hCAII immobilized on both spots were performed so that four single measurements on four independent hCAII surfaces could be taken on two consecutive days. Table 6-4 lists the results of the single measurements, the mean values and the corresponding standard deviations.

**Table 6-2 Results of the single measurements on four independent protein surfaces for ligands 54-58.**

| Ligand | Spot 1 (=3700 RU) | | | | Spot 2 (=3600 RU) | | | |
|---|---|---|---|---|---|---|---|---|
| | $K_D$ (st-st) [nM] | $K_D$ (kin) [nM] | $k_{on}$ [ x $10^5$ 1/Ms] | $k_{off}$ [1/s] | $K_D$ (st-st) [nM] | $K_D$ (kin) [nM] | $k_{on}$ [ x $10^5$ 1/Ms] | $k_{off}$ [1/s] |
| **54** | 1002.0 | 1073.4 | 1.6 | 0.1696 | 1037.0 | 1183.9 | 1.5 | 0.1828 |
| **55** | 382.2 | 318.8 | 1.9 | 0.0617 | 429.8 | 333.5 | 2.0 | 0.0664 |
| **56** | 133.7 | 91.4 | 5.1 | 0.0464 | 157.3 | 93.8 | 5.4 | 0.0504 |
| **57** | 63.4 | 57.4 | 8.7 | 0.0498 | 66.8 | 64.5 | 9.6 | 0.0617 |
| **58**[A] | 452.8 | 418.3 | 1.6 | 0.0673 | 488.6 | 448.0 | 1.7 | 0.0741 |
| **Ligand** | **Spot 1 (=4900 RU)** | | | | **Spot 2 (=4700 RU)** | | | |
| | $K_D$ (st-st) [nM] | $K_D$ kin) [nM] | $k_{on}$ [ x $10^5$ 1/Ms] | $k_{off}$ [1/s] | $K_D$ (st-st) [nM] | $K_D$ (kin) [nM] | $k_{on}$ [ x $10^5$ 1/Ms] | $k_{off}$ [1/s] |
| **54** | 1076.0 | 1063.1 | 1.4 | 0.1465 | 1426.0 | 1278.4 | 1.2 | 0.1529 |
| **55** | 218.0 | 228.8 | 2.2 | 0.0498 | 296.8 | 290.3 | 1.9 | 0.0550 |
| **56** | 104.3 | 100.3 | 4.3 | 0.0426 | 106.1 | 107.8 | 4.2 | 0.0451 |
| **57** | 53.1 | 51.4 | 6.7 | 0.0345 | 59.0 | 57.9 | 6.2 | 0.0357 |
| **58**[A] | 358.2 | 330.3 | 1.7 | 0.0567 | 414.8 | 371.6 | 1.6 | 0.0591 |

[A]For ligand **58** a sample with a purity of 94% was used. As explained in the experimental part the results do not significantly differ from the sample with 98% purity.

**Table 6-3 Results for the measurements of 58 with two different degrees of purity.**

| Ligand | 58 (94%) | | | | 58 (98%) | | | |
|---|---|---|---|---|---|---|---|---|
| | $K_D$ (st-st) [nM] | $K_D$ (kin) [nM] | $k_{on}$ [ x $10^5$ 1/Ms] | $k_{off}$ [1/s] | $K_D$ (st-st) [nM] | $K_D$ (kin) [nM] | $k_{on}$ [ x $10^5$ 1/Ms] | $k_{off}$ [1/s] |
| Spot 1(= 4900 RU) | 442.5 | 381.9 | 1.6 | 0.0599 | 404.2 | 359.2 | 1.7 | 0.0609 |
| Spot 2(= 4800 RU) | 476.7 | 416.1 | 1.6 | 0.0649 | 444.3 | 387.6 | 1.7 | 0.0657 |
| Spot 1 (=4500 RU) | 584.6 | 508.2 | 1.3 | 0.0665 | n.d. | n.d. | n.d. | n.d. |
| Spot 2 (=4500 RU) | 509.6 | 437.9 | 1.4 | 0.0602 | 414.1 | 394.2 | 1.5 | 0.0603 |
| Mean | 503.4 | 436.0 | 1.5 | 0.0629 | 420.9 | 380.3 | 1.6 | 0.0623 |
| Standard Deviation | 60.7 | 53.3 | 0.1 | 0.0034 | 20.9 | 18.6 | 0.1 | 0.0030 |

**Table 6-4 Validation against reference results.**

| Ligand | 59 (4-CBS) | | | | 60 (sulfanilamide) | | | |
|---|---|---|---|---|---|---|---|---|
| | $K_D$ (st-st)[B] [µM] | $K_D$ (kin)[B] [µM] | $k_{on}$ [x $10^4$ 1/Ms] | $k_{off}$ [1/s] | $K_D$ (st-st)[B] [µM] | $K_D$ (kin)[B] [µM] | $k_{on}$ [x $10^4$ 1/Ms] | $k_{off}$ [1/s] |
| Spot 1(= 4900 RU) | 0.99 | 0.89 | 6.5 | 0.0579 | 6.17 | 5.23 | 3.9 | 0.2023 |
| Spot 2(= 4800 RU) | 0.92 | 0.95 | 6.4 | 0.0605 | 7.17 | 5.82 | 3.7 | 0.2147 |
| Spot 1 (=4500 RU) | 1.01 | 0.90 | 6.2 | 0.0557 | 6.21 | 5.38 | 3.6 | 0.1963 |
| Spot 2 (=4500 RU) | 0.96 | 0.82 | 6.8 | 0.0559 | 6.24 | 5.77 | 3.6 | 0.2080 |
| Mean | 0.97 | 0.89 | 6.5 | 0.0575 | 6.45 | 5.55 | 3.7 | 0.2053 |
| Standard Deviation | 0.04 | 0.05 | 0.3 | 0.0022 | 0.48 | 0.29 | 0.1 | 0.0079 |
| Literature (bCAII) [A] | n.d. | 0.89 | 4.1 | 0.0369 | n.d. | 5.88 | 2.3 | 0.133 |

[A] Results for the measurements of the reference compounds 4-CBS and sulfanilamide with hCAII and comparison with data for bCAII from literature.[60]
[B] The definition of the steady-state and kinetic dissociation constants ($K_D$) are reported in Section 9.3.1.4.

Although we worked with human CAII, we were able to measure similar affinity data for **59** (hCAII: 0.89 µM, bCAII: 0.89 µM) and **60** (hCAII: 5.55 µM, bCAII: 5.88 µM) as described in the literature for the bovine type. Furthermore, the kinetic constants show a similar trend, identifying **59** ($k_{on}$ hCAII: $6.5x10^4$ $M^{-1}s^{-1}$; $k_{on}$ bCAII: $4.1x10^4$ $M^{-1}s^{-1}$; $k_{off}$ hCAII: 0.0575 $s^{-1}$; $k_{off}$ bCAII: 0.0369 $s^{-1}$) as the compound with the faster association and slower dissociation with respect to **60** ($k_{on}$ hCAII: $3.7x10^4$ $M^{-1}s^{-1}$; $k_{on}$ bCAII: $2.3x10^4$ $M^{-1}s^{-1}$; $k_{off}$ hCAII: 0.2053 $s^{-1}$; $k_{off}$ bCAII: 0.1330 $s^{-1}$). Additionally, this series of measurement showed that the assay was reproducible when similar amounts of protein were immobilized on the sensor chip.

The running buffer in an SPR experiment often contains 0.005-0.05% Tween 20 to prevent unspecific binding to the tubings and other hydrophobic parts of the instruments. To exclude the possibility that Tween 20 interferes with the binding of ligands **54**-**58** to hCAII, the same series of measurements were performed twice with the same protein surface, once with 0.05% Tween 20 in the running buffer and once without the detergent. The results are listed in Table 6-5.

**Table 6-5 Effects of Tween 20 addition.**

| | With 0.05% Tween 20[A] | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Spot 1 (=3700 RU) | | | | Spot 2 (=3600 RU) | | | |
| | $K_D$(st-st) [µM] | $K_D$ (kin) [µM] | $k_{on}$ [ x $10^5$ 1/Ms] | $k_{off}$ [1/s] | $K_D$ (st-st) [µM] | $K_D$ (kin) [µM] | $k_{on}$ [ x $10^5$ 1/Ms] | $k_{off}$ [1/s] |
| **54** | 1002,0 | 1073.4 | 1.6 | 0.1696 | 1037,0 | 1183.9 | 1.5 | 0.1828 |
| **55** | 382.2 | 318.8 | 1.9 | 0.0617 | 429.8 | 333.5 | 2.0 | 0.0664 |
| **56** | 133.7 | 91.4 | 5.1 | 0.0464 | 157.3 | 93.8 | 5.4 | 0.0504 |
| **57** | 63.4 | 57.4 | 8.7 | 0.0498 | 66.8 | 64.5 | 9.6 | 0.0617 |
| **58** | 452.8 | 418.3 | 1.6 | 0.0673 | 488.6 | 448.0 | 1.7 | 0.0741 |
| | **Without Tween 20[A]** | | | | | | | |
| | Spot 1 (=3700 RU) | | | | Spot 2 (=3600RU) | | | |
| | $K_D$(st-st) [µM] | $K_D$ (kin) [µM] | $k_{on}$ [ x $10^5$ 1/Ms] | $k_{off}$ [1/s] | $K_D$ (st-st) [µM] | $K_D$ (kin) [µM] | $k_{on}$ [ x $10^5$ 1/Ms] | $k_{off}$ [1/s] |
| **54** | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. | n.d. |
| **55** | 292.3 | 265.9 | 2.1 | 0.0554 | 312.2 | 277,0 | 2.0 | 0.0547 |
| **56** | 122,0 | 107.1 | 4.5 | 0.0483 | 126.4 | 113.2 | 4.6 | 0.0523 |
| **57** | 66.1 | 60.5 | 7.8 | 0.0472 | 66.1 | 62.8 | 8.4 | 0.0528 |
| **58** | 412.8 | 386.1 | 1.5 | 0.0596 | 425.2 | 402.2 | 1.7 | 0.0667 |

[A]The same series of measurements was perfomed two times on the same protein surface. One time 0.05% Tween 20 was added to the running buffer and one time omitted. Similar values were obtained in both cases.

Similar values were obtained for ligands **55**-**58** in both cases. Ligand **54** showed sensorgrams which did not reach saturation when Tween 20 was not added. Thus, we did not evaluate the data for **54** in this case. Since **54** showed a satisfying saturation when adding 0.05% Tween 20, the measurements were done in the presence of the tenside.

For the final measurements of the investigated ligand series (**54**-**58**), single measurements on four independent protein surfaces were performed. To partially average out the small variation of the kinetic constants depending on the immobilization level, we decided to determine the mean of these four surfaces exhibiting different protein densities. We performed experiments with immobilization levels at a lower rate corresponding to 1000 RU. However, the quality of the obtained sensorgrams was not sufficient. Our results are not significantly influenced by a mass

transfer limitation even at a high immobilization level. Thus, the applied immobilization levels were chosen since they led to a good signal-to-noise ratio.

The $k_{on}$ values determined for compounds **54**-**58** (Figure 6-1 and Figure 6-3b-c) by SPR measurements are consistent with previously reported SFF results. Although ligands **57** and **58** are of similar length, the less hydrophobic compound (**58**) displays a much smaller $k_{on}$. Conversely, a significant change in $k_{off}$ is experienced between **54** and **55** only, while **55**-**58** all possess very similar dissociation rate constants. We find that the trend for the equilibrium dissociation constants $K_D$ is driven by variation in the association kinetics, in line with the results obtained by King and Burgen.[66]

### 6.4.3 Results of MD simulations

Molecular dynamics (MD) simulations highlighted common hCAII-ligand interactions for the alkyl sulfonamide series (compounds **54**-**57**). In particular, we observed that the ligands approach the pocket entrance by mainly exposing the polar sulfonamide to the bulk solvent, while trying to segregate the alkyl tail in a partly exposed pocket formed by hydrophobic residues. A conformational analysis revealed the exact identity of the amino acids involved in the hydrophobic interaction, which we took as the residues having minimum distance from the ligand below 3 Å in the most populated conformational cluster. This group of residues defined a partially solvent-exposed patch (hereafter referred to as the "hydrophobic patch", HP), which grew as the size of the substituent increased (Figure 6-2 and Figure 6-3a). HP was largest for compound **57**,where it included amino acids Ile91, Val121, Phe131, Val135, Leu198, Pro201 and Pro202 (according to the numbering of the PDB code 1CA2),[172] and partly corresponded to the portion of hCAII known in the literature as the hCAII "secondary hydrophobic binding site"[173] (Figure 6-3a). After binding to the HP domain, the ligand remained transiently bound to the protein and, in most of the cases, was subsequently released back into the water bulk. A quantitative description of the release process was obtained by computing, for all ligands, the lifetime of the intermediate state at the HP domain ($t_{HP}$). The lifetimes were computed by observing the time evolution of the minimum distance between the ligand and the HP patch, this distance being typically characterized by abrupt increases, which corresponded to ligand detachments from the protein surface (see Section 6.6.1.5). The lifetimes of **54**-**57** increased with

the length of the alkyl chain as shown in Figure 6-3b-c. In some instances, the ligands turned the sulfonamide's polar head towards the catalytic site, reaching the hCAII catalytic pocket.

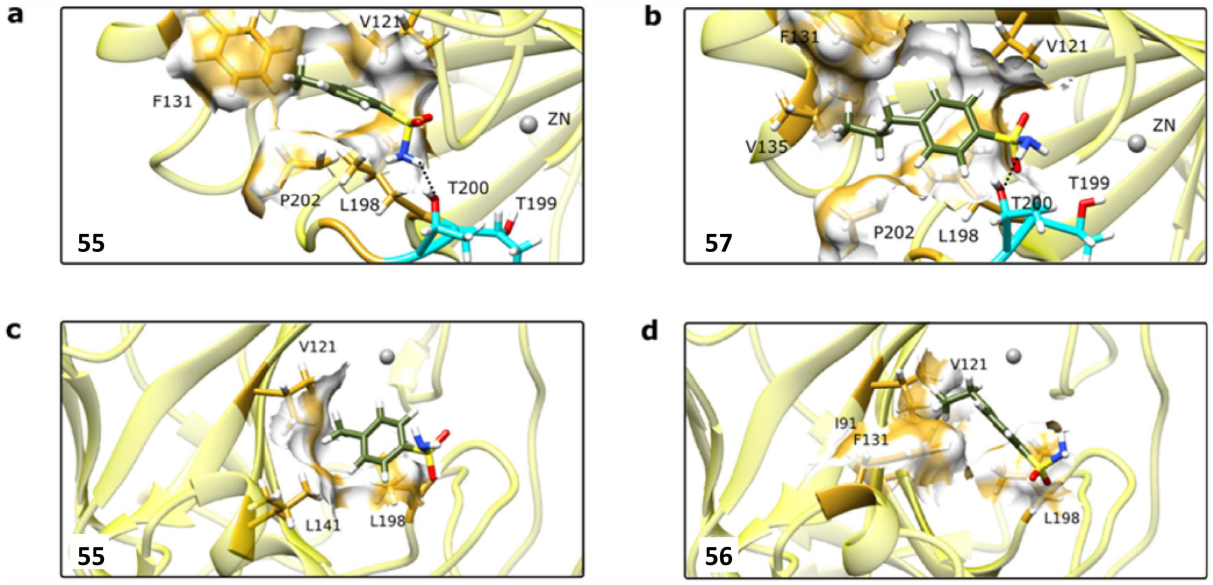
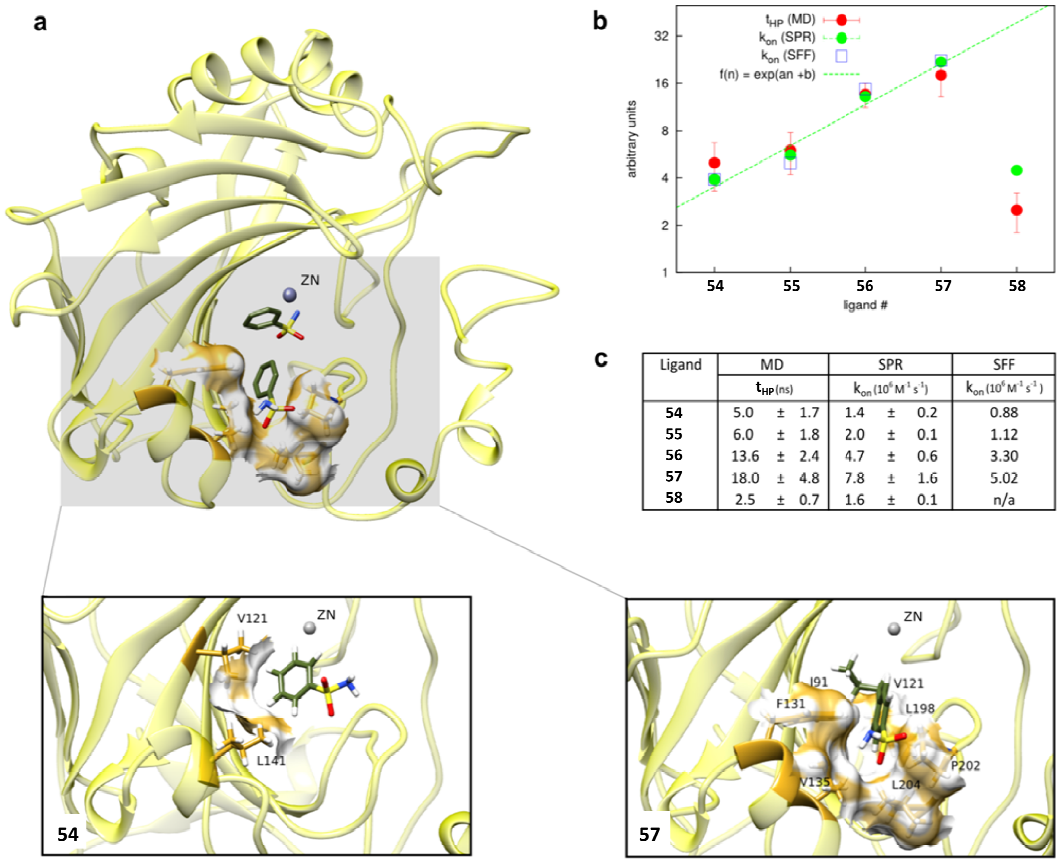



**Figure 6-2 Conformational clustering.** (a,b) Most populated S conformations found in plain MD for ligands **55** and **57**. (c,d) Most populated F conformations found in plain MD for ligands **55** and **56**.

We named these conformations "straight" (S), as opposed to the purely HP-bound geometries, which we labeled as "flipped" (F). Both S and F conformations could be partitioned into representative clusters (Figure 6-2, Figure 6-3a and Figure 6-4a). In one case (compound **54,** Figure 6-4a), the amide group of the sulfonamide was found to directly interact with the $Zn^{2+}$-bound water molecule, which is taken to be a hydroxyl ion in the present simulations. Hydrogen bond interactions between the ligands **54**-**57**, in the S conformations and the backbone oxygen of Thr199 and Thr200 played a key stabilizing role. In addition, in the most populated S conformations, the alkyl tails of benzenesulfonamides formed favorable hydrophobic packing with the apolar side of the hCAII cavity. The largest hydrophobic cage was observed for ligand **57** and it was composed of residues Val121, Phe131, Val135, Leu198 and Pro202 (Figure 6-2b). A completely different scenario was observed with compound **58**. In the latter case, the lifetime of the ligand on the HP patch was the shortest within the entire series (**54-58**, Figure 6-3b-c), due to unfavorable interactions between the hydrophilic tail and HP. No buried S conformations were

observed for ligand **58** during MD simulations. We could hypothesize that the transition from F to S is energetically too costly to be extensively sampled by plain MD simulation. For this reason, we resorted to so-called enhanced sampling methods (see below).




| Ligand | MD | SPR | SFF |
|---|---|---|---|
| | $t_{HP}$ (ns) | $k_{on}$ ($10^6$ $M^{-1}$ $s^{-1}$) | $k_{on}$ ($10^6$ $M^{-1}$ $s^{-1}$) |
| 54 | 5.0 ± 1.7 | 1.4 ± 0.2 | 0.88 |
| 55 | 6.0 ± 1.8 | 2.0 ± 0.1 | 1.12 |
| 56 | 13.6 ± 2.4 | 4.7 ± 0.6 | 3.30 |
| 57 | 18.0 ± 4.8 | 7.8 ± 1.6 | 5.02 |
| 58 | 2.5 ± 0.7 | 1.6 ± 0.1 | n/a |

**Figure 6-3 Structure-kinetic relationships**. (a) Structural overview. The clipped semi-transparent surface represents the secondary hydrophobic region at the entrance. The F conformation of the unsubstituted ligand is exemplified together with the $Zn^{2+}$-bound geometry, for comparison. In the inset, we represent the explicit F conformations of ligands **54**, **57** and report the hydrophobic residues with which a contact is established. (b) Comparison between the pre binding lifetime, $t_{HP}$ (see text), obtained in MD and the $k_{on}$ measured by SFF and SPR experiments. Data are reported in logarithmic scale and arbitrary units on the y-axis. The green dotted line represents a linear fit on the SFF data (see text). (c) Table containing the actual $t_{HP}$ and $k_{on}$ values, reported with their standard deviations.

### 6.4.3.1 Structural characterization of the pre-binding stage

The conformational sampling was enhanced by means of metadynamics simulations within the hCAII binding cavity. We focused on the compounds displaying the most sizable structural differences, i.e. the smallest and largest ligand of the alkyl-substituted series (**54** and **57**), as well as the hydroxyl-terminated compound **58**. The free energy profile was projected over two collective variables, namely CV1 (distance between $Zn^{2+}$ and the sulfonamide sulfur) and CV2 (angle formed by the $Zn^{2+}$ atom, the sulfonamide sulfur and the carbon atom bound to the sulfonamide sulfur), devised to accelerate the translation (CV1) and rotation (CV2) of the sulfonamides within the cavity (see also Figure 6-4a). Mainly two free energy minima were found, corresponding to the S and F conformations observed in MD simulations. For all compounds, the lowest free energy minimum corresponded to S conformations characterized by a CV1 value of about 6 Å and CV2 ranging from 90° to 180°. In the energetic minimum, we found a largely visited conformation, in which the sulfonamide interacted with the $Zn^{2+}$-bound hydroxyl ion, showing an average distance between the sulfonamide nitrogen ($N_{sulf}$) and the $Zn^{2+}$ of 3.8 ± 0.3 Å. This geometry was the one with the smallest $Zn^{2+}$-$N_{sulf}$ distance obtained during our simulations and was characterized by two hydrogen bonds between the amide group of sulfonamides and both the Thr200 side chain oxygen and the $Zn^{2+}$-bound hydroxyl. An additional hydrogen bond was observed between the sulfonamide oxygen and the Thr199 side chain. A pictorial representation of the S conformation is reported in Figure 6-4. For ligands **54**, **57** and **58** we observed a secondary free energy minimum, corresponding to F conformations and well represented by the conformational clusters reported in Figure 6-3 and in Figure 6-5. These geometries were stabilized by a favorable hydrophobic interaction between the ligand chain and the HP patch. As for compound **58**, the free energy of the F conformation, relative to the global minimum S state, was much higher than in the case of **57**. We estimated an energy separation between the F and the S basins for **54**, **57** and **58** of roughly 9, 12 and 22 kJ/mol, respectively (5 kJ/mol being the error estimated on convergence, see Section 6.6.1.3). In the same order, the energy of saddle points which divide the F from the S region was about 10, 11 and 8 kJ/mol higher than the energy of the F conformation, reflecting the energetic barrier associated with the rotation of the compound during the F to S transition. The overall free energy landscape is reported in Figure 6-4. Compound **58** showed an additional low free energy region, which

extended approximately over the range of CV1 values between 15 and 25 Å. In the corresponding conformations, both favorable hydrophobic packing and hydrogen bonding contributed to the interaction between the ligand and the residues at the protein surface. As an explanatory example, the structure corresponding to the minimum energy of this external region is provided in Figure 6-5. Both the stability of the F conformation and the presence of the additional low energy region are expected to impact on the overall association kinetics of **58**.

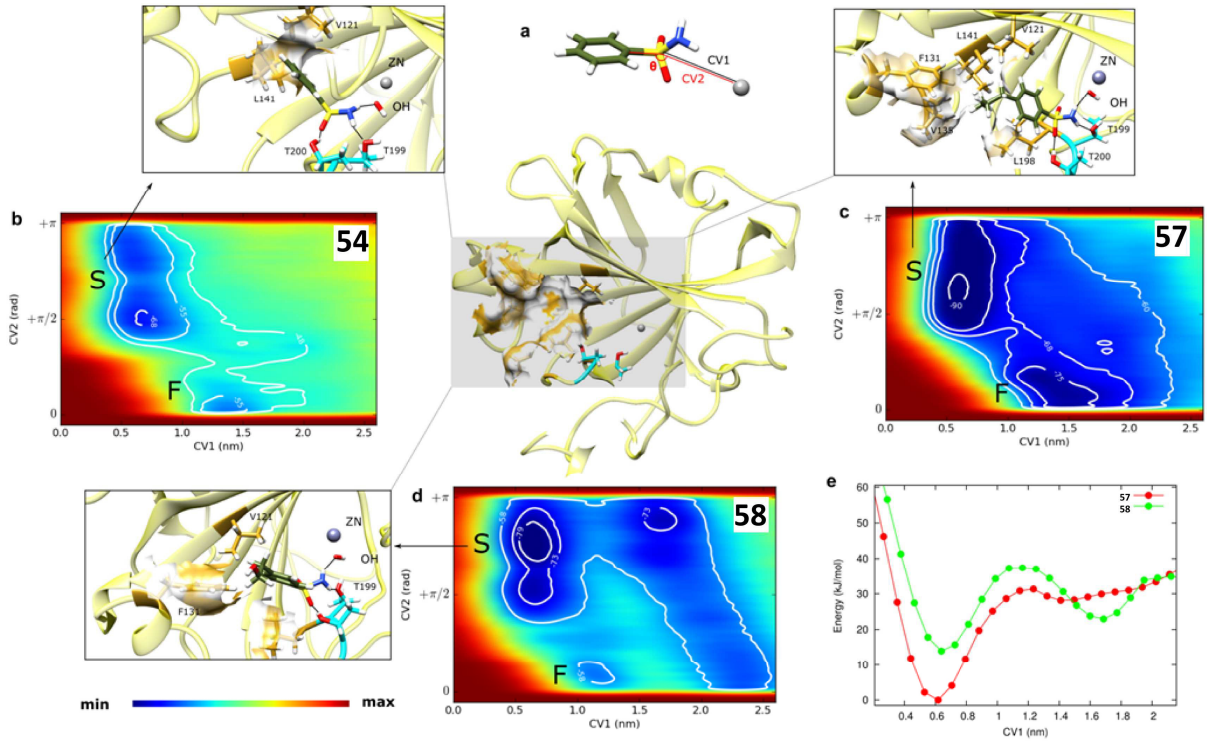



**Figure 6-4 Free energy landscape of the protein-ligand complexes.** (a) Description of CV1 and CV2 and structural overview of the protein. (b,c,d) Free energy profiles and conformations associated with the free energy minima for compounds **54** (b), **57** (c) and **58** (d). In the structural description reported in inset, the ligands and the residues within 3 Å of the S conformations are shown in stick representation. Hydrophobic and hydrophilic residues have a dark yellow and cyan skeleton, respectively. H, N, O and S atoms are respectively white, blue, red and yellow. In the free energy profiles, contours with the associated energy value in kJ/mol are reported. (e) Projection of the free energy profile on the CV1 for ligands **57** and **58** showing the different barriers associated with the two compounds.

In Figure 6-4e, we report the projection of the free energy profiles of **57** and **58** on CV1. In this projection, the barrier associated with the entrance of the sulfonamide in the cavity was averaged over the rotational degrees of freedom. The hydroxyl-substituted ligand showed a substantially

larger barrier than the methyl-substituted compound, pointing to the penalty due to the desolvation of **58**.

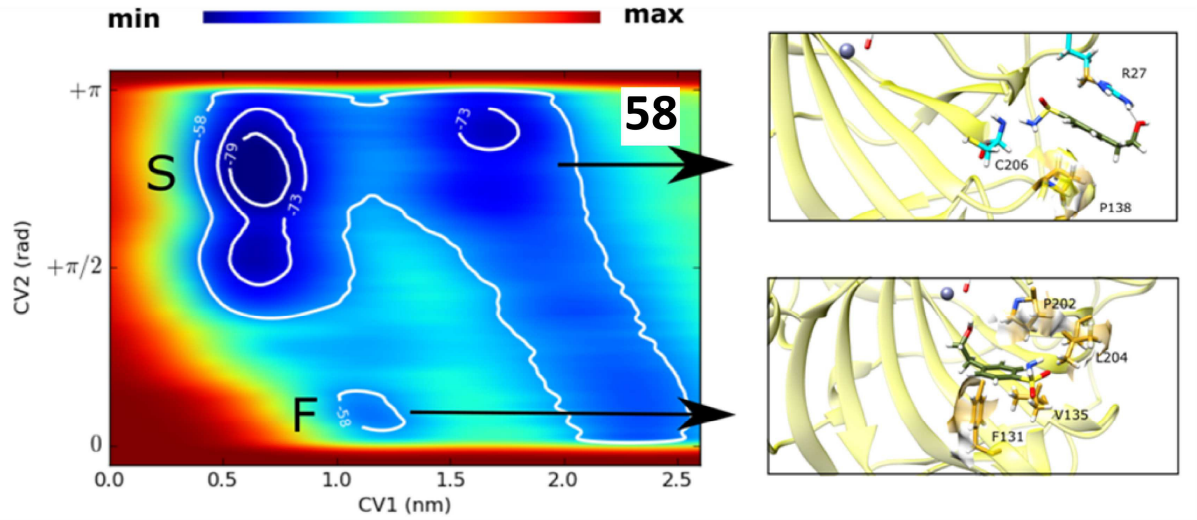



**Figure 6-5 Conformations in the secondary free energy minima of ligand 58.** Figure explanation follows the caption of Figure 6-4.

### 6.4.3.2 Analysis of sulfonamide acidity

The sulfonamide moiety changes its protonation state upon binding, but the configuration where the proton-transfer mechanism occurs has not yet been identified. Therefore, we used a continuum electrostatics-based computational approach to evaluate the propensity of the sulfonamide to deprotonate in the F structures encountered during dynamics. The approach, based on the Poisson-Boltzmann equation, yielded for the ligand **57** a pKa shift of 0.3 between the F conformation and the fully solvated structure. However, DFT-based geometry optimization performed on the S conformation led to a geometry in which a neutral sulfonamide forms a hydrogen bond with the $Zn^{2+}$-bound hydroxyl molecule. The propensity was not reversed even when the initial conformation of the system corresponded to a deprotonated sulfonamide facing a $Zn^{2+}$-bound water molecule. The $Zn^{2+}$-$N_{sulf}$ distance at which the sulfonamide transfers one proton to the hydroxyl molecule was estimated by performing a series of DFT-based constrained geometry optimizations, in which the $Zn^{2+}$-$N_{sulf}$ distance was restrained at a set of target values. The sulfonamide deprotonated spontaneously at a distance between 2.50 and 2.75 Å (see Figure 6-6 for further details).

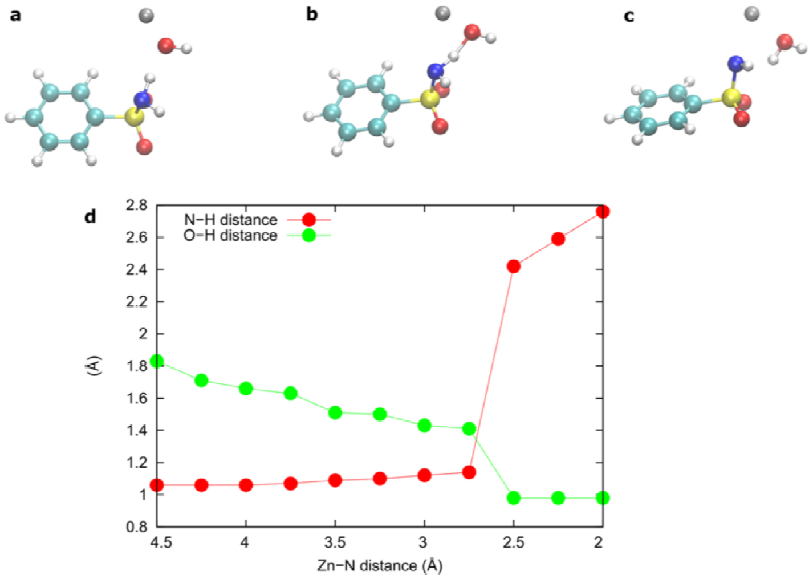



**Figure 6-6 Sulfonamide-hydroxyl proton transfer:** (a,b,c) Subset of the system used for the constrained geometry optimizations. The benzenesulfonamide, the hydroxyl ion (or water molecule) and the zinc ion are reported. The images depict the optimized geometry at a Zn-N distance of 3.50 (a), 2.75 (b) and 2.50 Å (c).(d) The length of the N-H bond in the sulfonamide, as well as the O-H distance between the hydroxyl oxygen and the incoming proton are reported as a function of the Zn-N distance.

### 6.4.4 End of the association process: crystallographic structures

The full characterization of the association process requires the knowledge of the final state, which cannot be achieved using classical MD simulations. The stable $Zn^{2+}$-bound conformation of the sulfonamides was instead obtained experimentally by X-ray crystallography. The geometry of the ligands in their crystallographic structure offers precious insights into the structure-kinetics relationship and the role of water networks within the protein cavity. The crystal structures of ligands **54-58** in complex with hCAII (Figure 6-7-Figure 6-9) were determined at high resolution (hCAII-**54**: 1.01 Å, hCAII-**55**: 0.96 Å, hCAII-**56**: 1.06 Å, hCAII-**57**: 1.08 Å, hCAII-**58**: 1.07 Å). The sulfonamide group forms with one of its oxygen atoms the typical interactions with the amide nitrogen (2.9 Å) and the side chain oxygen atom (2.8 Å) of Thr199.[174, 175] Furthermore, the nitrogen atom of the sulfonamide coordinates to the catalytic zinc ion (2.0 Å). In Figure 6-7a, we report the alignment of the common scaffold of **54**-**58**. The structure of hCAII-**54** (1.45 Å,

PDB: 2WEJ) was previously published, but at a lower resolution.[174] Our structure is in good agreement with the former one, to which it adds more details, such as two additional water molecules (W7 and W8, Figure 6-8).

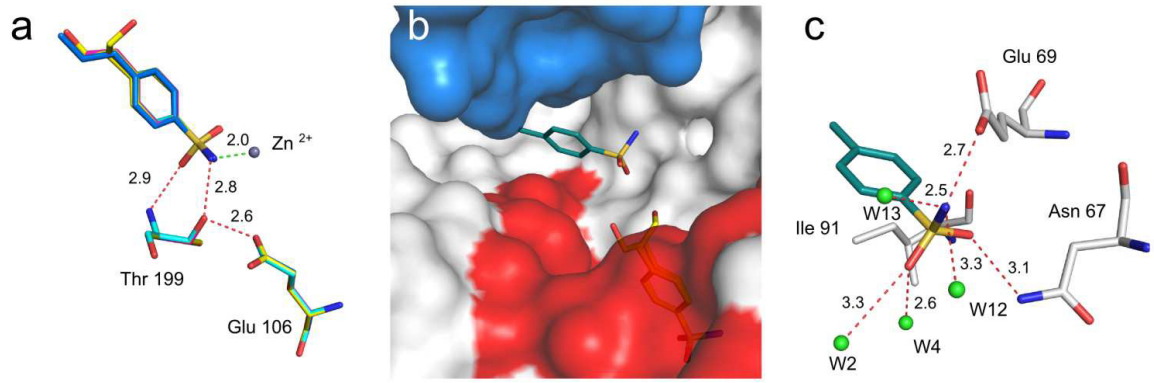



**Figure 6-7** a) Alignment of crystal structures hCAII-**54** – hCAII-**58**. Oxygen atoms are colored in red, sulfur atoms in yellow and nitrogen atoms in blue. The zinc ion is shown as a grey sphere. H-bond distances are shown as red-dotted lines. The distance between the zinc ion and the nitrogen of the ligands is shown as a green-dotted line. This color code is maintained throughout the following figures unless stated elsewise. Carbon atoms of the ligand and of the corresponding protein residues are colored in green for **54**, in orange for **55**, in purple for **56**, in light blue for **57** and in yellow for **58**. The binding modes of **54**-**58** align nearly perfectly. b) The binding positions of **58** in the active site and at the surface next to the hydrophobic patch are shown. The surface of hCAII is colored in grey, the residues which form the hydrophobic patch are colored in red. The surface of the hCAII molecule of the next symmetry mate is colored in blue. One molecule of **58** binds as expected deeply buried in the active site (carbon atoms are colored in yellow). A second molecule of **58** binds additionally at the entrance of the binding pocket with contacts to the hydrophobic patch (carbon atoms are colored in turquoise). Only the first carbon atom of the hydroxyethylene tail of **58** was modeled into the density since the remaining atoms were not well defined in the difference electron density. c) Interactions of the second molecule of **58**. Carbon atoms of the protein residues of the hCAII are shown in grey. Water molecules are shown as green spheres. Van der Waals contacts to Ile91 which is part of the hydrophobic patch are detectable. These figures and the following representations were prepared using PyMOL.[36]

We collected the structural data using 25% (v/v) glycerol which is routinely used as a cryoprotectant for hCAII crystals. Aggarwal et al. showed that glycerol is able to displace three water molecules.[176] An overlay of our hCAII-**54** complex structure with a high resolution *apo*-structure of hCAII (PDB: 3KS3,[53] 0.90 Å ) shows a similar finding (Figure 6-9a). The waters shown as blue- and as red-colored spheres were found in the *apo*-structure, those waters shown as green colored spheres belong to the hCAII-**54** structure. In the hCAII-**54** structure the glycerol displaces the three water molecules shown in blue which are detected in the *apo*-structure. Interestingly, waters W2-W5 of the hCAII-**54** complex are found in equal position in the *apo*-

structure. So we estimate that the glycerol molecule does not influence the water structure significantly.

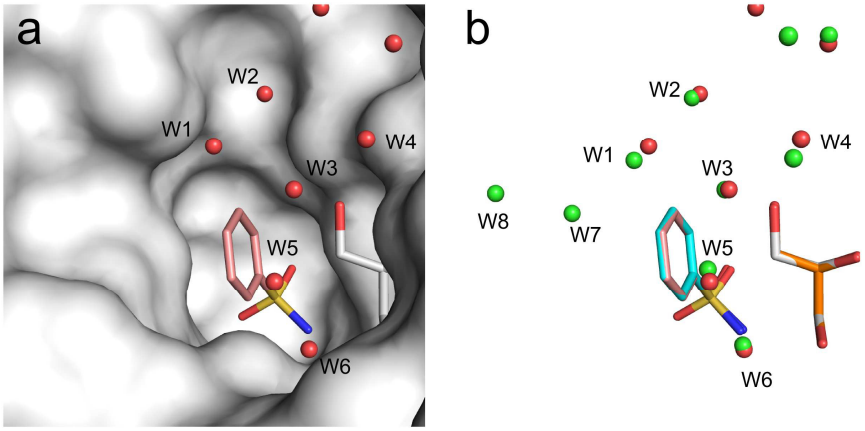



**Figure 6-8** a) Previously published crystal structure of **54** complexed to hCAII (1.45 Å, PDB: 2WEJ[174]). Oxygen atoms are colored in red, sulfur atoms in yellow and nitrogen atoms in blue. Carbon atoms of **54** are colored in salmon. This color code is maintained throughout the following figures unless stated elsewise. b) Alignment of the hCAII-**54** structure which is reported in this work (1.01 Å) with the structure with the PDB code 2WEJ. The carbon atoms of **54** of the here reported hCAII-**54** structure are shown in cyan. The positions of **54**, glycerol and the water molecules W1-W6 align very well. Two additional waters (W7, W8) can be placed in the difference density in our structure. Presumably they were not detected in the previous structure due to the lower resolution.

The waters W2-W6 are found in very similar positions for ligands **54-58** (Figure 6-9b-f). For ligands **54**-**57** the Lee-Richard surfaces are shown. The Lee-Richard surface is an estimation of the solvent accessible surface area (SASA).[177] All detected waters coincide or fall close to the surface of this SASA. W1 is present in the complexes hCAII-**55-58** at virtually the same position. In the hCAII-**54** structure it can penetrate deeper into the pocket since **54** does not possess a hydrophobic tail pushing W1 to the position which it occupies in hCAII-**55**-**58**. In hCAII-**54** W1 constructs an H-bond network with waters W7 and W8. Water W7 is only present in the hCAII-**54** structure because only here it can be stabilized by an H-bond with W1. Furthermore, in hCAII-**55**-**57** it will be displaced owing to the steric requirements of the hydrophobic tail of the ligand. Only for **54** the position of W7 falls outside the Lee-Richard surface around the ligand. W8 is also visible in the hCAII-**56** and hCAII-**58** structures. In hCAII-**55** some $F_o$-$F_c$ difference electron density (3.13 σ, 0.50 e/Å$^3$) could be detected at the position for W8 (Figure 6-9c) but its

shape and magnitude did not fully justify the placement of a full water molecule in the refinement. However, W8 seems to be stabilized in the hCAII-**55** complex, but not to the same extent as in hCAII-**54**, hCAII-**56** and hCAII-**58**. A density peak suggests occupancy of W8 in hCAII-**56**, however, without being connected to the remaining active site water molecules (Figure 6-9d). An additional water (W11) is visible in the hCAII-**56** structure. It lies above W1 and W3 and is stabilized by H-bonds to the latter water molecules. Both W8 (3.4 Å) and W11 (4.0 Å) are stabilized by a van der Waals contact to the terminal methyl group of **56**. Similar contacts are visible between the ethyl group of **55** to W1 (3.6 Å) and the terminal methyl group of **57** to W1 (3.5 Å). In hCAII-**57** no evidence for the presence of W8 could be detected. This is in good agreement with the Lee-Richard surface (Figure 6-9b-e) which extends as a consequence of the additional carbon atom of **57** into the region being occupied by W8 in hCAII-**54**, -**56** and -**58**. The terminal hydroxyethylene group of **58** is visible in two orientations. The higher populated orientation A (58%) forms an H-bond with the glycerol molecule while the orientation B (42%) is similarly orientated as the tail of **57**. In orientation B the terminal OH group of **58** forms H-bonds to the neighboring water molecules W1 and W8. We assume that orientation B is the predominant one in the glycerol-free condition since the described H-bond between the terminal OH group of the tail of **58** in the orientation A to the glycerol could not be established. Orientation A will exclude presence of W3 as the OH group of **58** (1.7 Å) would come very close. Accordingly, W3 is only present when **58** is in orientation B. Ligand **54** is the only one which is bordered by a contiguously connected water structure on the side towards the solvent. This ligand shows a significantly faster dissociation than **55**-**58**. Based on the crystallographic analysis the deviating water structure is the most striking difference of **54** compared to **55**-**58**, thus possibly the better established water network favors the faster dissociation of **54** since the ligand might be easier displaced by water molecules.

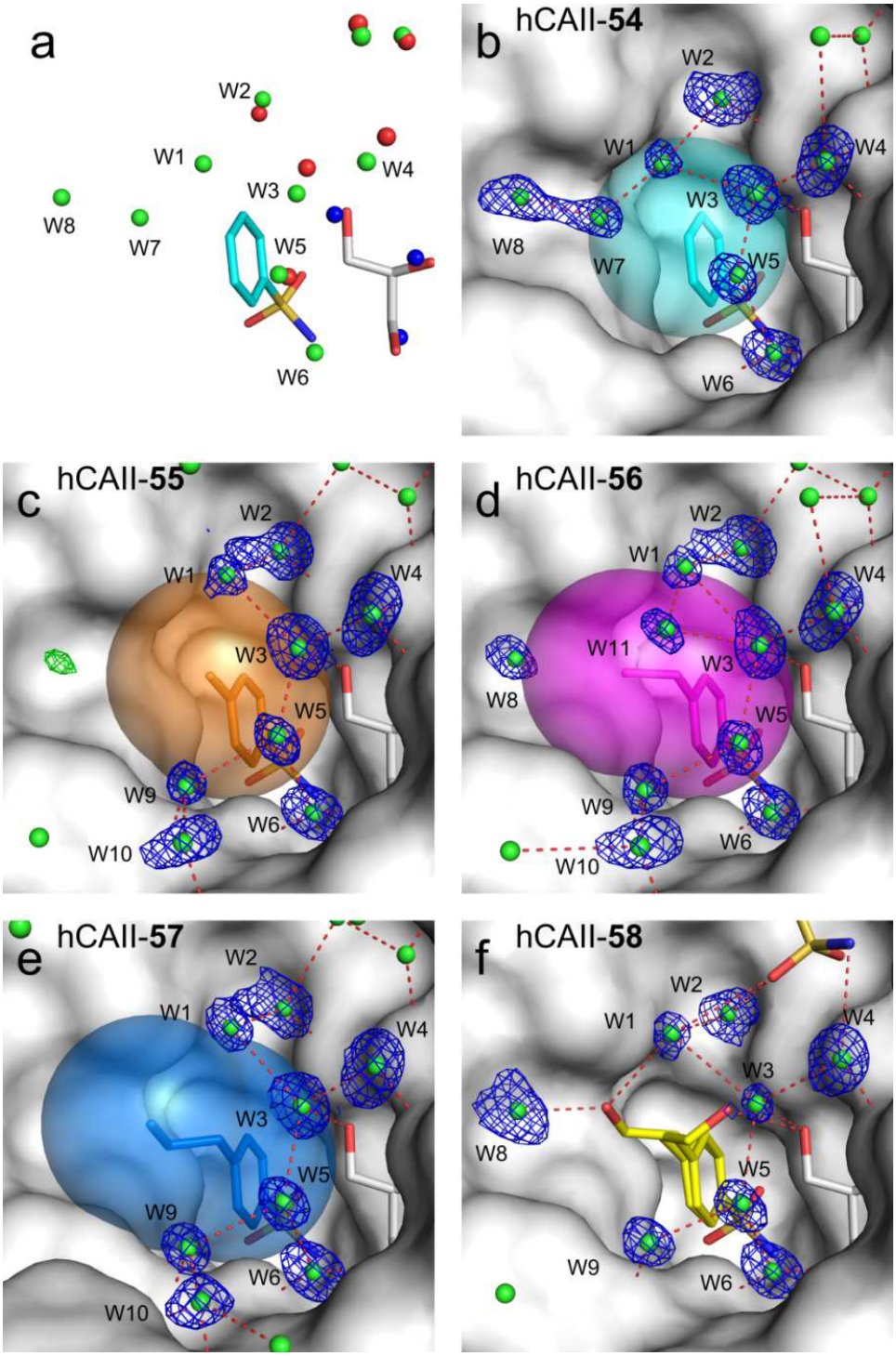



**Figure 6-9.** Detailed figure caption on next page.

**Figure 6-9 Crystallographic Structures.** a) Comparison of the water structures in complexes of the hCAII-**54** and of the *apo*-protein (PDB: 3KS3[53]). The water molecules in hCAII-**54** are shown as green spheres. The carbon atoms of the glycerol molecule which is present in all hCAII complexes are shown in grey, the carbon atoms of **54** are shown in cyan. The water molecules of the *apo*-structure which are displaced by the glycerol molecule are shown as blue spheres, the water molecules of the *apo*-structure which are also present in the hCAII-**54** structure are shown in red. b)-f) Binding mode of **54** (b), **55** (c), **56** (d), **57** (e) and **58** (f) in complex with hCAII. The protein surface is shown in gray, the carbon atoms of **54**, **55**, **56**, **57** and **58** are colored in cyan, orange, magenta, blue and yellow, respectively. The carbon atoms of glycerol are colored in gray. Water molecules are shown as green spheres. The $2F_o$-$F_c$ electron density is contoured in blue at 0.50 electrons/Å$^3$ (0.9 σ for **54**, **55, 57** and **58** and 1.0 σ for **56**) for water molecules W1-W8. H-bond distances (2.3- 3.3 Å) are depicted as red-dotted lines. The Lee-Richard surfaces for **54**-**57** are shown in the same color as the carbon atoms of the respective ligands. They were generated for the hydrophobic tails of the ligands, starting with the 4-carbon atom of the phenyl ring. A contiguously connected water network, which borders the ligand at the solvent-exposed side, is detectable in the hCAII-**54** structure only. The tail of **58** is visible in two orientations.

However, as these changes occur at the already solvent-exposed face of the ligands it remains unclear how strongly the water structure influences the dissociation velocity. For the complex structures of ligands **55**-**58** additional molecules of the respective ligand could be found in the $F_o$-$F_c$ difference electron densities. For **55**-**57** one additional ligand, for ligand **58** three additional ligands could be placed. Crystal packing effects together with the high concentrations of the ligands applied in the soaking buffer might be responsible for this observation. The additional binding position observed for **58** adjacent to the active site (Figure 6-7b-c) might suggest a putative position intermediately occupied by the ligands before they slip into the final binding position in the active site. This additional binding mode was only observed for **58**. Further investigation is needed to fully unravel the significance of this mode at lower ligand concentrations, since no conformation of this type was observed during MD simulations of one single fully solvated ligand-protein complex. Due to the additional polar group **58** shows a higher solubility than all other compounds. This might explain why we find only for this ligand three additional molecules in the $F_o$-$F_c$ difference electron density since saturated solutions of the ligands were applied. The terminal hydroxyethylene tail of this extra molecule of **58** could only be modeled up to the first carbon atom since the remaining atoms are not unambiguously defined in the difference electron density. Figure 6-7b shows the surface of the next symmetry mate in the crystal packing (colored in blue). It contributes to the stabilization of the binding mode of this additional ligand since it forms a cavity with the neighboring protein molecule (coloured in gray). Apart from the contact to the crystal mate also interactions to residues of the same protein molecule are formed thus suggesting that also the binding in solution could be established

(Figure 6-7b-c). The nitrogen of the additional molecule of **58** forms an H-bond to one of the oxygens of the carboxylic acid group of Glu69 (2.7 Å) while one of its oxygen atoms forms an H-bond with the carboxamide nitrogen of Asn67 (3.1 Å). Van der Waals interactions between the aromatic ring (4.1 Å) and the first carbon atom of the tail of **58** (4.5 Å) to the side chain of Ile91 are detected. Since this residue is part of the hydrophobic patch, it could be envisaged that this pose, similarly to the ones located at the secondary patch, is also stabilized by favorable hydrophobic interactions. Additional binding configurations, stabilized by a subtle interplay of protein-ligand interactions and crystallographic packing are presented in Figure 6-10**.**

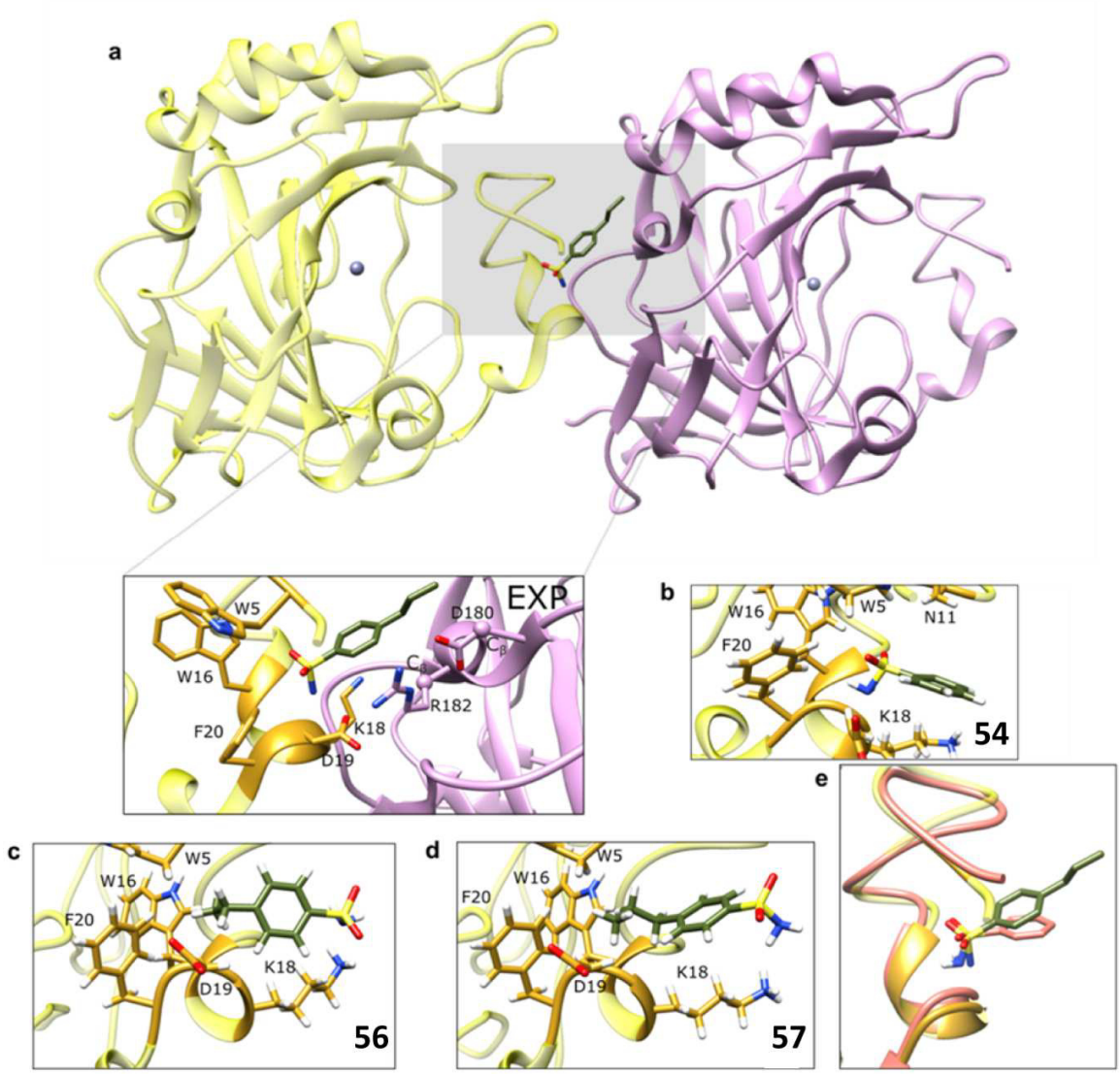



**Figure 6-10 Binding modes at the N-terminal cleft.** (a) Structural overview of the crystallographic protein-ligand system. The ligand **57** is shown together with the two proteins (colored yellow and violet, respectively) with which it forms contacts. In the inset we show the residues whose distance from the ligand is less than 4 Å. The ligand forms interactions with the N-terminus cleft of the first protein (yellow) and the $C_\beta$ of Asp180 and Arg182 (minimum $C_\beta$-C) of the second protein (violet). The minimum distance between the $C_\beta$ and the sulfonamide carbons is 4.1 Å (Arg182) and 4.0 Å (Asp180). (b,c,d) Center of conformational clusters obtained by simulations. (b) ligand **54** (population = 0.6%) (c) ligand **56** (population = 5%) (d) ligand **57** (population = 6%) (e) Alignment at the N-terminus between the crystallography structure and the configuration reported in (b). The structures of ligand **57** and **58** in the N-terminal cleft were found predominantly with the polar sulfonamide pointing outwards, at variance with the experimental result. The crystallographic structure could not be recapitulated presumably because of the unfavorable interaction between the larger alkyl tail and the surrounding water. This interaction is on the other hand lifted in the crystal structure by the protein packing, thereby providing an explanation for the observed behavior.

## 6.5 Discussion and Conclusions

In agreement with previous SFF data, our SPR measurements show that the association rate for ligands **54**-**57** is dependent on their chemical composition. As already discussed by Taylor, King and Burgen,[66, 67] the structural dependence of $k_{on}$ does suggest the presence of at least one intermediate state. A single-step binding process would instead be compatible with a purely diffusion-limited binding phenomenon (about $10^8$ $M^{-1}s^{-1}$ in the case of ligand-enzyme association),[178] which has not been observed for the hCAII-sulfonamides complexes considered here. The comparison of SPR data for compound **57** and the newly analyzed compound **58** provides additional information on the nature of the binding intermediates. Compound **57** displays a considerably faster association than **58** and, since these two molecules differ only by the replacement of a methyl with a hydroxyl group in the chain, this indicates that hydrophobicity is key for the binding kinetics to this enzyme. MD simulations support this view and show how the favorable hydrophobic packing helps to precisely stabilize the geometries along a multi-step binding pathway. Starting from the solvated state in the bulk solution, the compounds approach the protein by ligating to an outer secondary hydrophobic site, adopting a particular conformation, named here "F conformation". As shown in Figure 6-3, hydrophobic packing is the only interaction that makes F conformations largely populated. Similarly, at a later stage of the binding process, the ligands adopt what we named the "S conformation", in which the apolar moieties of the compound's substituents still largely interact with the hydrophobic patch of hCAII (Figure 6-2 and Figure 6-4), while the sulfonamide group is already engaged in hydrogen bonding with the protein and the $Zn^{2+}$-bound hydroxyl ion. The S geometry may be regarded as the final conformation adopted by the ligand, before coordination with the $Zn^{2+}$ ion occurs. DFT-based calculations show that the sulfonamide deprotonation can occur when the distance between the sulfonamide nitrogen and $Zn^{2+}$ falls between 2.50 and 2.75 Å, leading finally to the displacement of the hydroxyl ion and the formation of the coordinative bond between the deprotonated sulfonamide and the enzyme's zinc ion. The structures of these final states between the sulfonamides **54**-**58** and hCAII were obtained by means of X-ray crystallography and are reported in Figure 6-9.

A comment is here required on the deprotonation mechanism of sulfonamides, which has remained an open issue in the literature.[46] It might be argued that sulfonamides change their protonation state before reaching the S conformation. There are, however, good reasons to consider this an unlikely hypothesis. First, in the pre-binding F conformation, the pKa value of the sulfonamide changes negligibly (0.3, according to our pKa-shift estimate), which makes a deprotonation event in the F conformation very unlikely. Secondly, the approach to the catalytic $Zn^{2+}$ as observed in the S conformation, does not make the sulfonamide acidic enough for a proton transfer. Indeed, DFT geometry optimization of S predicts that the proton transfer from the sulfonamide to the hydroxyl ion cannot occur spontaneously, requiring a shorter distance between the sulfonamide group and the $Zn^{2+}$. Therefore, the present computational study points to a mechanism in which the sulfonamides approach hCAII in their neutral state, bind to the hydrophobic patch and interact with the $Zn^{2+}$. Then, only if the distance between their terminal $NH_2$ group and the $Zn^{2+}$ is about 2.50-2.75 Å, will they get deprotonated and coordinate the $Zn^{2+}$ ion, resulting in the final thermodynamic minimum.

Overall, our analysis, at least for the alkyl-substituted derivatives (compounds **54**-**57**), suggests a three-step process, regulated by partial kinetic constants as described in Figure 6-11 and in the scheme below:

$$[P][L] \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} [F] \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} [S] \overset{k_3}{\longrightarrow} [PL] \qquad (6-1)$$

where [P][L] represents the unbound protein-ligand system, [F] and [S] correspond to the previously discussed conformations and [PL] stands for the $Zn^{2+}$-bound sulfonamide complex. We can get an insight into the dependence between the partial kinetic constants and the observed $k_{on}$ by first considering the conversion between [P][L] and [F] as a rapid pre-equilibration step. This is justified by the analysis of our MD simulations, showing that $k_{-1} \gg k_2$.

In the pre-equilibration hypothesis the unbound system is taken to be in equilibrium with the F basin and yields:

$$k_1[P][L] = k_{-1}[F] \qquad (6-2)$$

We now work under the hypothesis that the variation of [S] during the reaction is small enough for steady state kinetics to be applied:

$$k_2[F] - k_{-2}[S] - k_3[S] = \frac{\partial[S]}{\partial t} = 0 \qquad (6-3)$$

The rate of formation of the bound state at the last association step is:

$$k_3[S] = \frac{\partial[PL]}{\partial t} \qquad (6-4)$$

Combining equation (6-2) and (6-3) it is possible to express [S] in terms of the unbound state concentration [P][L]. Plugging the expression into (6-4) we get:

$$\frac{k_1}{k_{-1}}\frac{k_2 k_3}{k_{-2}+k_3}[P][L] = \frac{\partial[PL]}{\partial t} = k_{on}[P][L] \quad (6-5)$$

If we assume the change in concentration of [S] will be small, i.e. we apply steady-state kinetics for the second and third step, we reveal the following overall association constant (see (6-2)-(6-5) for a simple derivation) as:

$$k_{on} = (k_1/k_{-1}) \cdot (k_2 k_3/(k_{-2}+k_3) \qquad (6-6)$$

In the derivation we set $k_{-3} = 0$, similarly to what done by King and Burgen[66, 67] for the dissociation step in the simpler two-state mechanism.

The rate associated with the ligand diffusion, $k_1$, as well as the deprotonation rate, $k_3$, may be considered weakly dependent on the chemical decoration throughout the ligand series **54**-**57**. Actually, $k_1$ is determined by the Brownian diffusion of the ligand in solution and only weakly depends on the hydrodynamic radius of the molecule, which remains nearly constant across the series **54**-**57**. Similarly, changes in the substituents within the ligands investigated occur rather

remote from the sulfonamide group and are not expected to largely influence the amide group acidity. Actually, in bulk water, the pKa difference between **54** and **57** is about 0.2.[66, 67] Moreover, the barriers for conversion between the F and S conformations are predicted to be similar (see Figure 6-4), at least for the alkyl-substituted ligands. Inspired by these considerations, we conjecture that the $k_{on}$ trend along **54**-**57** should mainly correlate with the inverse of the partial unbinding rate $k_{-1}$, which can be estimated by the lifetimes of the ligands at the hydrophobic patch (HP). The explicit calculation of $t_{HP}$ (see Figure 6-3) supports our hypothesis. The release kinetics, described by $k_{-1}$, of the intermediate complex F also provides an

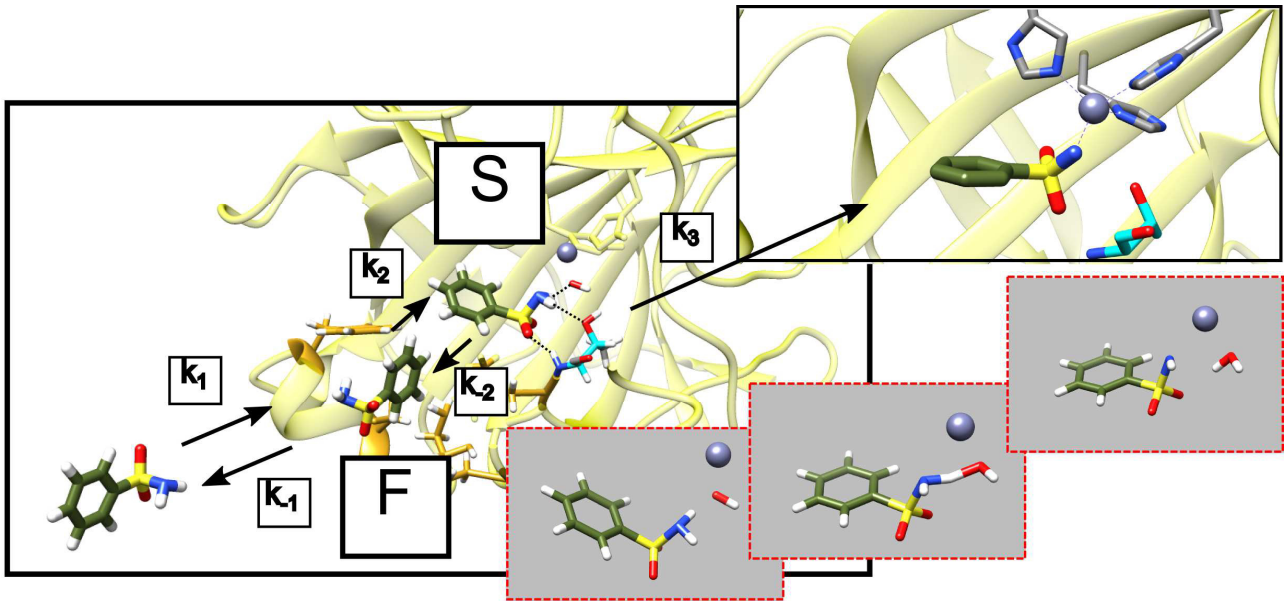



**Figure 6-11 Overview of the binding process:** The unbound ligand approaches the protein and reaches the F conformation, located in the external hydrophobic patch, with rate $k_1$. At this stage, the ligand might return to the unbound state with rate $k_{-1}$ or reach the S conformation with rate $k_2$. From the S state, characterized by geometries displaying hydrogen bonding with T199, T200 and the $Zn^{2+}$-hydroxyl ion contact, the compound can possibly leave the site with rate $k_{-2}$ or reach the final crystallographic structure (see inset) with rate $k_3$. Final isomerization to the $Zn^{2+}$-bound pose requires deprotonation of the sulfonamide moiety and conversion of the $Zn^{2+}$-bound hydroxyl ion to a water molecule. The proton transfer mechanism is sketched in the sequence of bottom right insets.

explanation for the exponential dependence of $k_{on}$ on the alkyl chain length, for the ligand series **54**-**57**. Adding one methylene group to the alkyl chain energetically stabilizes the F structures by an amount, $\Delta G_{CH2}$, which represents the free energy penalty associated with the transfer of one

methylene group from the protein's apolar environment to the water solution. By adopting a simple Arrhenius relationship, the rate $k_{-1}$ decreases by a factor exp ($-\beta\Delta G_{CH2}$) when a methylene group is added to the substituent chain. The fit of the Arrhenius expression to the SFF data leads to a value of $\Delta G_{CH2}$ of 1.56 ± 0.25 kJ/mol (see Figure 6-3b). The free energy penalty, per $CH_2$ group, associated with the transfer of an alkyl chain from a lipophilic to aqueous solvent was experimentally determined to be 3.35 kJ/mol.[147, 179] This value can be considered consistent with our estimation of $\Delta G_{CH2}$, as the benzenesulfonamides desorb from a hydrophobic surface (rather than from a bulk hydrophobic solution), which was in contact with approximately one half of the total volume occupied by the alkyl chain.

The experimental $k_{off}$ values, obtained by SPR experiments, are nearly independent of the compound structures, in agreement with previous SFF results.[66, 67] The notable exception is ligand **54**, the $k_{off}$ of which is nearly three times bigger than that measured for the other compounds. The crystal structure of **54** in complex with hCAII shows a unique arrangement of crystallographic waters compared to the other derivative, which fully surround the ligand and connect it to the solvent exposed side. This could indicate the crucial role of a water network in the observed $k_{off}$ value. On the other hand, hydrophobicity does not play a role in determining the off rates: the presence of the hydroxyl group in **58** does not lead to a dramatic reduction in the $k_{off}$ with respect to the similarly sized ligand **57**. In summary, the affinity of benzenesulfonamides for the hCAII target appears to be mostly modulated by changes in the association rate. This study explains the trends observed, on the basis of a structure-kinetics relationship retrieved by the analysis of large-scale atomistic simulations. hCAII is an excellent system for investigating the basic mechanisms determining the affinity of an enzyme for a specific inhibitor. In this case, we have shown that hydrophobicity determines the rate at which a class of benzenesulfonamides binds to hCAII. Moreover, this study demonstrates that computational modeling can be used to reconstruct the complete binding pathway of a ligand to its biological target, even in the highly non-trivial case where an early slow recognition process must be described together with a proton transfer reaction. We believe that our work offers a paradigmatic approach, showing how the combination of several experimental and simulative

methods can be generally used to produce a detailed understanding of the enzyme-inhibitor complexation mechanisms.

## 6.6 Description of Experimental Procedures

### 6.6.1 Computational section

#### 6.6.1.1 Protein preparation

hCAII coordinates were retrieved from the Protein Data Bank (PDB: 1CA2).[172] Amber99sb[180] force field was used for modeling the standard residues. The catalytic site was described by the Zinc Amber Force Field (ZAFF) force field.[181] The crystallographic water bound to the zinc ion was assumed to be a hydroxyl ion.

#### 6.6.1.2 Ligand preparation

Ligands **54**-**58** (Figure 6-1) were parameterized using the Generalized Amber Force Field (GAFF).[182] Charges were refitted by means of restrained electrostatic potential (RESP).[183] The three lowest energy structures, obtained by *in-vacuo* conformational search using Macromodel[184] and employing the Optimized Potential for Liquid Systems' (OPLS),[185 186] were used for charge fitting.

#### 6.6.1.3 Molecular dynamics and metadynamics simulations

MD simulations were run for 1 μs (**54**) 1.2 μs (**55**) 1.1 μs (**56**) 1.1 μs (**57**) 1 μs (**58**) using the Gromacs 4.5.5 package.[187] Simulations were run in the NPT ensemble, using the Parrinello-Rahman barostat[188] with time constant $\tau$ = 2 ps and the V-rescale thermostat[189] with time constant $\tau$ = 0.1 ps. A time step dt = 2 fs was used. Geometry optimization was performed on the initial conformation up to a force threshold of 10 kJ $mol^{-1}$ $Å^{-1}$ and the resulting structure was brought from 0 to 300 K using a constant annealing of 400 ps. Well-tempered metadynamics[190] simulations were performed on **54**, **57** and **58** using the PLUMED plugin.[191] The duration of each run was at least 150 ns per ligand. We used two collective variables, CV1 and CV2 defined as the Zn-S1 distance and the Zn-S1-C1 angle, S1 and C1 representing, respectively, the sulfur of the sulfonamide and the carbon bound to it (see also Figure 6-4a). The initial Gaussian height was set to 2.5 kJ $mol^{-1}$, while Gaussian widths were set to 2.5 Å and 0.5 rad for CV1 and CV2, respectively. The Gaussian deposition time interval was set to 2 ps. Simulations were stopped

after at least 5 recrossings from the lowest free energy conformation. This allowed the difference between the minima of the free energy profiles to remain constant within 5 kJ $mol^{-1}$ over the final 25 ns of the simulation.

#### 6.6.1.4 Definition of the simulation region

We defined a simulation region as an ensemble of spheres with a 5Å radius, centered on each atom of the following residue group: Trp5, Trp16, Phe20, Pro21, I22, Ile91, Val121, Trp123, Phe131, Val135, Leu141, Ala142, Val143, Leu198, Pro201, Pro202, Leu203, Leu204 and Trp209. We considered the following minimum distance collective variable:

$$s = \frac{\beta}{log\sum_{ij} exp\left(\beta/\| r_{ij}\|\right)} \quad (6\text{-}7)$$

Where $\|r_{ij}\|$ represents the distance between the *i-th* atom of the ligand and the *j-th* atom of the residue group. β was set to 500.

The constraining potential was introduced as:

$$U(s) = 0 for s < s_0 \quad (6\text{-}8)$$

$$U(s) = k(s - s_0)^4 for s \geq s_0 \quad (6\text{-}9)$$

where $s_0$ = 5 Å and k= $4 \cdot 10^2$ kJ $mol^{-1} \cdot Å^{-1}$. Using these conditions, the ligand, when lying completely outside of the confining region, experiences a restraining force, preventing the escape of the ligand itself from the confinement volume. This allows sampling of a larger number of detachment events with respect to unconfined simulations, without biasing the dynamics during the event of ligand desorption. The shape of the constraining region as it appears in the hCAII crystal structure is displayed in Figure 6-12. Because of the relatively large Gaussian width employed, metadynamics simulations were allowed to sample a larger volume, i.e. the spheres used to define the simulation region were set to 10 Å.

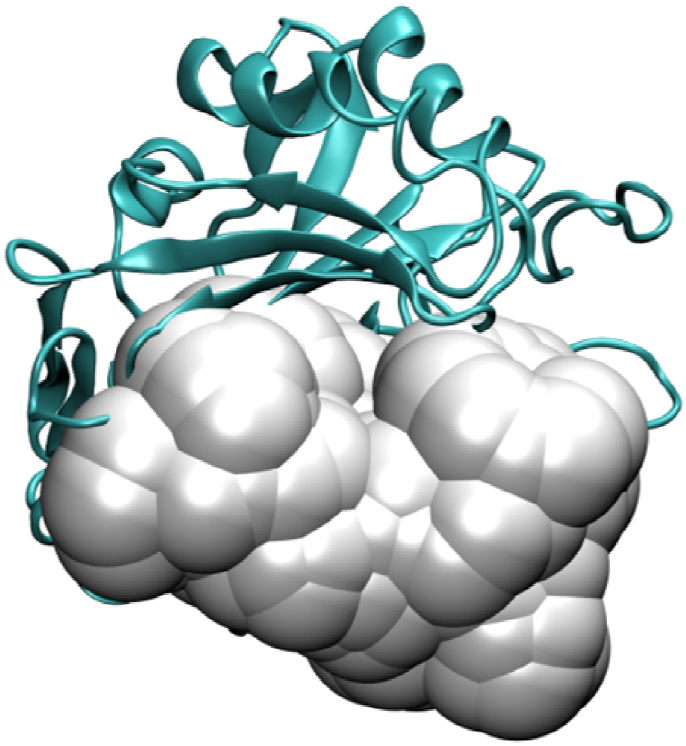


**Figure 6-12 Confining region of ligands in simulations.** Confining region, represented as a union of spheres (white balloons), overlaid on a cartoon representation of hCAII.

#### 6.6.1.5 Pre-binding lifetime

The individual lifetimes, $t_{HP}$, of ligands on a predefined hydrophobic patch were computed by monitoring the time evolution of the minimum distance *s* between the ligand atoms and all the atoms of the patch. The considered patch need not be the whole restraining region defined above, but could be a portion of it, as specified in the main text (see definition of HP patch). In our calculations, the actual region used for lifetime calculations was the union of the HP patches for the ligands **54**-**57**, corresponding to a pocket comprising Ile91, Val121, Phe131, Val135, Leu141, Leu198, Pro201 and Pro202. This choice allowed us to focus on the lifetime of the relevant pre-binding conformations, which are stabilized by hydrophobic interactions and are precisely hosted in the above-mentioned patch. Two threshold values, $s_{low}$ and $s_{high}$, were introduced and used to tag the state as "adsorbed" (A) or "desorbed" (D) following four classification rules:

1) At time t=0, State(0)=D if s(0)> $s_{low}$ or State(0)=A if s(0)< $s_{low}$.

2) At any time $t_0$ such that: State($t_0$-dt)=A and s($t_0$)>$s_{high}$, then: State($t_0$)=D.

3) At any time $t_0$ such that: State($t_0$-dt)=D and s($t_0$)< $s_{low}$, then: State($t_0$)=A.

4) In all other cases State($t_0$)=State($t_0$-dt).

In our calculations, we used $s_{low}$=2 Å and $s_{high}$=5 Å. The first value can be considered the distance at which a hydrophobic contact between the ligand and the HP patch is actually established. The second value corresponds to a rather minimal distance at which the ligand and the patch can be separated by a water molecule. The total length of an adsorption event defines the individual lifetime, $t_{HP}$. At least 20 desorption processes per ligand have always been observed. The procedure to compute $t_{HP}$ is exemplified in Figure 6-13. The distributions of the lifetimes are reported for all ligands in Figure 6-14. Conformational cluster analysis was performed using the Gromos method.[192] The RMSD cutoff radius employed was 2 Å and atoms included in the RMSD calculation were the *Cα* of the residues in the restraining patch and the ligands' atoms.

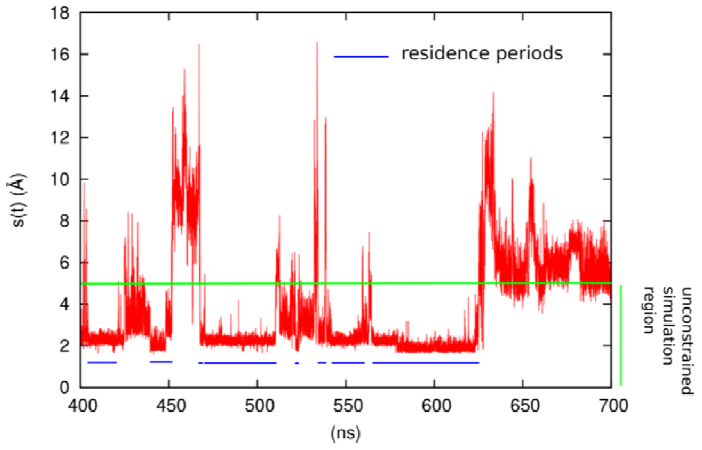



**Figure 6-13 Example of $t_{HP}$ determination.** The minimum distance s(t) between the ligand and the HP patch (see Section 6.6.1.5) is plotted during part of the trajectory of hCAII-**56** system. The broken blue line identifies the different residence periods, which are delimited by abrupt spikes of s(t). For values of s(t) larger than the limit depicted by the horizontal green line, a constraining potential favors the repositioning of the ligand on the HP patch, thus enhancing the number of binding/unbinding events in the simulation.

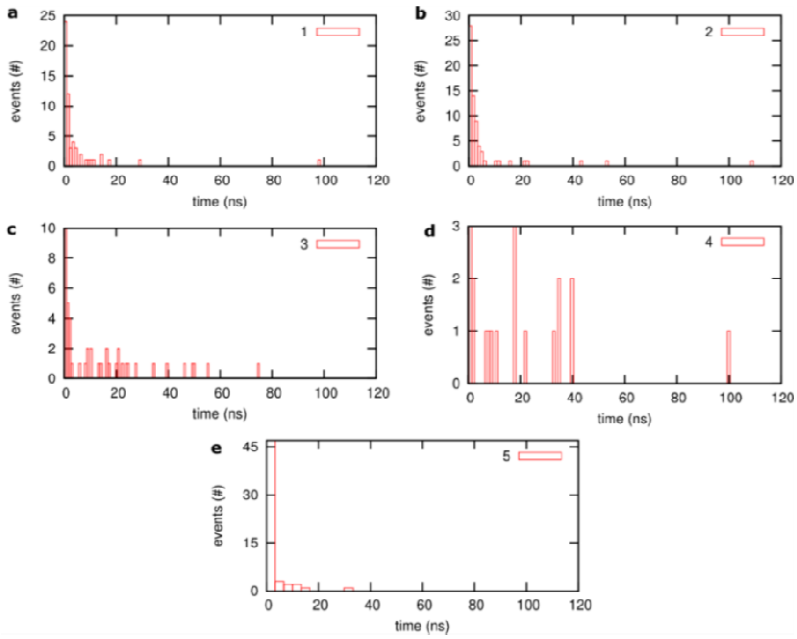



**Figure 6-14 Distribution of the $t_{HP}$ in the HP pocket. 54** (a), **55** (b), **56** (c), **57** (d), **58** (e).

### 6.6.1.6 Poisson-Boltzmann calculations

A thermodynamic cycle was considered, connecting the protonated and deprotonated sulfonamide molecules in aqueous solvent, where the pKa is known experimentally and the same compounds once in the presence of the hCAII protein at a desired location. Considering the overall low charge of the system, the linearized Poisson-Boltzmann equation was solved for each of the four systems by means of the finite differences method provided by the DelPhi v.4 code.[193] The grid spacing was set to 0.5 Å and particular care was devoted to maintaining the same relative position of the sulfonamides with the finite differences grid in all the simulations. Coulombic-type boundary conditions and a percentage filling of 70% were used. The pKa shift was obtained by considering the difference between the electrostatic energy of protonated and deprotonated forms in the presence and absence of the protein.

### 6.6.1.7 DFT calculations

Calculations were performed on two systems (1 and 2) constructed as follows. (1): the S conformation was taken from dynamics for ligand **54** and optimized up to a force threshold of 10 kJ $mol^{-1}$ $Å^{-1}$. We then selected the residues coordinating the ligand, resulting in a system

comprising His94, His96, His119, His64, Glu106, Thr199 and Thr200. In the reduced model used, the residues have been cut at $C_\beta$ and one extra hydrogen was added to complete the coordination of the $C_\beta$. All waters found within 4 Å of the ligand were included in the calculation. (2): system (1) was modified by translating the H atom involved in the H-bond between the sulfonamide and the $Zn^{2+}$-bound hydroxyl ion along the H-O bonding distance. The resulting conformation corresponded to a $Zn^{2+}$-bound water molecule facing the deprotonated sulfonamide. Both systems were structurally optimized up to a force threshold of 4.5 $10^{-4}$ atomic units. In (1) and (2), all the $C_\beta$ and the oxygen atoms of the water molecules with truncated coordination were kept fixed during DFT geometry optimization. In (2), the geometry was optimized by first fixing the sulfonamide nitrogen and the transferred hydrogen and by then relaxing the whole system. All electron DFT calculations were performed using the CRYSTAL14 codes. Basis sets were taken from the crystal website library (http://www.crystal.unito.it/basis-sets.php).

Constrained geometry optimization was performed by applying an external harmonic force between the sulfonamide nitrogen and the $Zn^{2+}$ ion. The spring stiffness was set to 0.5 atomic units. Starting from the DFT-optimized S pose, the equilibrium length of the spring was gradually brought to 2.0 Å in ten steps. Each step corresponded to a new geometry optimization, with input coordinates taken from the system optimized at the previous iteration. At each step, the equilibrium length was reduced by 0.25 Å. The CP2K code[194] was used for the constrained geometry optimizations. The force threshold for structural optimization was set to $4.5 \cdot 10^{-4}$ atomic units. Goedecker-Teter-Hutter pseudopotentials[195] were used for all elements. All the electrons in the highest energy electronic shell were considered as valence electrons. In the case of Zn, the semicore 3d electrons were also included. Double zeta valence basis sets with one polarization function[196] were used.

### 6.6.2 Experimental section

#### 6.6.2.1 Expression, purification, crystallization, crystal structure determination and SPR measurements of hCAII

These methods are described in Section 9.3.

#### 6.6.2.2 Ligand verification and purification

##### 6.6.2.2.1 Ligand verification

The ligands **54**, **55**, **59** and **60** were purchased from Sigma Aldrich. **56** was purchased from Apollo Scientific Ltd., **57** and **58** from Ukrorgsyntez Ltd. (UORSY). Melting points were determined on a Leitz HM-Lux melting point microscope and are uncorrected. Proton NMR spectra were recorded on a JEOL ECA-500 MHz spectrometer ($^1$H NMR: 500.2 MHz) or a JEOL ECX-400 MHz spectrometer ($^1$H NMR: 399.8 MHz). Quantitative NMR (qNMR) measurements[197, 198] regarding compound **58** were exclusively performed at 500.2 MHz. As Certified Reference Material (CRM), maleic acid (δ 6.26 ppm), purchased from Sigma Aldrich, was used as internal reference standard. Chemical shifts are stated in parts per million (ppm) and were referenced to tetramethylsilane (TMS) at δ 0.00 ppm, except for qNMR measurements, where the solvent residual peak for DMSO-$d_6$ at δ 2.50 ppm served as reference. qNMR purities for the different purity levels of compound **58** were obtained from the calculation of the observed integral values of the CRM and the aromatic 2`,6` protons in relation to their molar proportion to each other using the formula described below.[197] NMR instrument parameters were in accordance with those required for standard qNMR measurements.[198] The NMR data were processed employing Delta NMR Processing and Control Software, version 5.0.0. Abbreviations concerning signal multiplicity are as follows: bs = broad singlet, d = doublet, dm = doublet of multiplets, m = multiplet, psq = pseudoquartet, q = quartet, s = singlet, sext = sextet, t = triplet. ESI+ mass spectra were obtained from a Q-trap 2000 triple quadrupol mass spectrometer (Applied Biosystems). Elemental combustion analyses were determined on a vario Micro cube CHNS analyzer (Elementar Analysensysteme GmbH). Chromatography was performed using prepacked silica gel columns (Interchim PuriFlash SI-HP, 30 μm). TLC was carried out using 0.2 mm aluminium plates coated with silica gel 60 $F_{254}$ (Macherey-Nagel) and the respective

substance spots were visualized by UV detection. Solvents for chromatography of compound **58** were used in analytical quality without further purification.

The analytical data is listed in Appendix 10.4.

#### 6.6.2.2.2 Ligand purification

Calculation of the purity (**P**) of the ligand **58** according to qNMR measurements was performed by the following formula[197]:

$$P = \frac{I_{analyte}}{I_{CRM}} x \frac{N_{CRM}}{N_{analyte}} x \frac{M_{analyte}}{M_{CRM}} x \frac{m_{CRM}}{m_{sample}} \quad x\, P_{CRM} \quad (6\text{-}10)$$

The calculations of the purity of **58** of the commercial substance, after the first and after the second purification step are shown in Appendix 10.5.

# 7 Conclusions and Outlook

The aim of drug discovery is the tailored design of ligands with desired properties for the respective target protein. Since the development of a new drug costs up to 1.6 billion dollars the decision, which of the screening hits should be pursued, is a critical step. To choose the right compound, knowledge about the complex recognition process between ligand and protein is important. The thermodynamic parameters, especially the enthalpic portion, might suggest which of the screening hits forms the strongest interactions without having access to the complex crystal structure. This concept is based on the hypothesis that enthalpic binders are better optimization candidates as on late stage optimization drug molecules are improved in the binding mainly for entropic reasons. The kinetic data give information about the association and dissociation behavior of a protein-ligand complex under consideration. Depending on the intended application, a ligand with a well-tailored residence time in the binding pocket appears promising for the desired therapeutic success. In this work, the binding of congeneric ligand series to their respective target proteins, ALR2 and hCAII, were analyzed by X-ray crystallography, ITC and SPR measurements in order to gain knowledge for further drug design cycles on these proteins but also to understand principles which can be used for related drug discovery projects.

For ALR2 several projects could be accomplished successfully. ALR2 is an interesting target for the prevention of diabetic complications. However, besides epalrestat, which is marketed as a drug in Japan and India, the strong efforts to find an effective and safe inhibitor of this protein were so far not successful. This work contributes with the investigation of compound **13** as a new type of ALR2 inhibitor. It originates from a database search and subsequent affinity and selectivity determination were performed by Drs. Milan Stefek, Marta Soltesova Prnova and Magdalena Majekova from the Slovak Academy of Sciences in Bratislava, Slovakia. The compound was found to be selective for ALR2 compared to two related proteins, aldehyde reductase and AKR1B10, which is necessary to prevent undesired side effects. In this work the highly-resolved complex crystal structure of **13** could be determined (1.26 Å) which led to two surprising findings. First, the specificity pocket of ALR2 was found in its closed state. This result is astonishing since the addressing of this pocket is assumed to be a prerequisite to design a

selective ALR2 inhibitor. Secondly, an additional molecule of this inhibitor was found more at the entrance of the pocket, within a distance which allows to form an H-bond to the first molecule of **13**. The question arose if this second molecule could always bind together with the first one. To answer this question ITC experiments were performed. They indicated that **13** binds with the same stoichiometry as other ALR2 inhibitors. In their corresponding complex crystal structures only one molecule could be found. Additionally, the decreased occupancy of 70% for the second molecule in the complex crystal structure suggests that its binding results from a rather high concentration of **13** which was applied in the soaking experiment. Summarized, **13** can be regarded due to its high affinity and selectivity for ALR2 as the first member of a new group of ALR2 inhibitors. The affinity of this ligand group might be further improved since the specificity pocket of ALR2 was not addressed. Other derivatives should be synthesized and the crystal structure complexed with ALR2 of the compounds with the highest affinities should be elucidated.

In a cooperation with the groups of Prof. Alberto Podjarny (CNRS, Illkirch, France), Dr. Ángel R. de Lera (Universidade de Vigo), Drs. Maria Jesús Martin and Jesús Ángel de la Fuente (Biomar Microbial Technologies S. A., León), Dres. Xavier Parés and Jaume Farrés (Universitat Autònoma de Barcelona) two compounds were investigated which inhibit both ALR2 and the related protein AKR1B10. This work contributed to this project with the determination of ITC data. In due course of binding to ALR2, a proton entrapment of 0.6 protons per mole of formed complex was found, which is in good agreement with previous results. Surprisingly, for the binding to AKR1B10, only an entrapment of 0.1 protons per mole of formed complex was found. This indicated that the protonation state in AKR1B10 prior to the binding of an inhibitor is different for AKR1B10 than for ALR2. In a following project, a study should be performed which clarifies this different behavior.

Furthermore, an ALR2 inhibitor series bearing a carboxylic acid group along with two or three aromatic systems was investigated using X-ray crystallography, $IC_{50}$ value determination and ITC. This group of ALR2 inhibitors derives initially from a virtual screening campaign. In this work the compounds resulting from an optimization cycle should be analyzed. The ligands were synthesized by the group of Prof. Martin Schlitzer (Philipps-University Marburg). Unfortunately,

using the cocrystallization or soaking condition in citrate buffer, which were prevalent at the beginning of this project, the ligands did not bind to ALR2. A new soaking condition in TRIS buffer was established to enhance the solubility of the inhibitors resulting in complex crystal structures for at least some of the ligands. It could be shown that all members of this group, for which a complex crystal structure could be determined, adopt a very similar binding mode. Differences could be found in the position of the carboxylic acid group in the anion binding pocket. Most members form a direct H-bond to the NE1 of Trp111 while two ligands establish a different interaction pattern. The latter two ligands capture a water molecule which compensates for the lost H-bond to Trp111. The first complex crystal structure of a ligand of this series possessing a third aromatic moiety could be obtained. It shows that the third ring is not hampering the binding in a similar geometry as for the other members of this series. Additionally, the complex crystal structure of the first ligand of this group possessing an acrylic acid head group was determined. It resulted in the surprising finding that the configuration of the incorporated double bond changed from *E* to *Z* during the soaking procedure. This leads to the conclusion that the *Z*-configuration is better able to form the interactions in the anion binding pocket than the *E*-configuration.

In another project the specificity pocket of ALR2 was investigated using derivatives of the already described inhibitor IDD393 which were synthesized by the group of Prof. Wibke Diederich (Philipps University of Marburg). Its portion which addresses the latter pocket was varied which led to derivatives which induced different conformations of the residues forming the specificity pocket. The comparison of two ligands, one which opens the pocket with a non-substituted phenyl ring and one ligand which has no moiety which could open the latter pocket, indicated that the opening of the pocket is accompanied by a small but measurable loss in the enthalpic portion. Since the complex crystal structure determination provided no explanation for this finding it can be assumed that the residues of the specificity pocket in the open conformation possess a higher energetic state compared to the closed conformation. Accordingly, to open the pocket, some energy has to be invested, leading to a small loss in the enthalpic portion. Different moieties were used to address the specificity pocket. A nitro group attached in *meta* position to the phenyl ring which binds in the specificity pocket leads to a surprisingly strong enthalpic gain

of -35.7 kJ/mol. It was shown previously that a nitro group at this position leads to a strong enthalpic signal, but in this work its contribution to the enthalpic binding signal could be quantified for the first time. The exchange of this nitro group by a carboxylic acid group surprisingly leads to a changed binding mode of the ligand. The terminal phenyl ring of the benzoic acid moiety is not buried in the specificity pocket as found for the nitro phenyl derivative but is directed to the solvent-accessible entrance of the binding pocket. To establish this binding mode, an internal H-bond between the amide nitrogen and one of the oxygen atoms of the carboxylic acid group of the inhibitor, which all other members of this ligand series establish, cannot be formed anylonger. This altered binding mode led to a 2000-fold decreased affinity. Presumably, the desolvation penalty for a carboxylic acid group is higher than for a nitro group, which could explain this surprising finding. In a following project different terminal moieties for this ligand group should be investigated to elucidate which is the discriminating factor determining whether the specificity pocket is open or not. This knowledge might be transferred to similar drug design projects which involve the addressing of a transient subpocket by a moiety of a ligand.

The binding of a series of benzenesulfonamide ligands to hCAII was characterized by SPR experiments and complex crystal structure determinations. The obtained results were compared to findings of MD simulations and previously published SFF measurements. Similar to the SFF experiments it could be shown that the association rates increase with the length of the hydrophobic tail which is attached in *para* position to the benzenesulfonamide. This is a surprising finding since the association rate of molecules with a long alkyl rest, at least in case binding would be only diffusion-controlled, should be hampered due to steric hinderence and consequently slowed down. Consequently, in this investigated situation, the association process seems to be a multi-step mechanism in which one step is favorably influenced by a hydrophobic interaction. The complex crystal structure determination showed that all investigated ligands exhibit a very similar binding mode. However, crystal structures show only the final binding position. With MD experiments, performed by the group of Prof. Andrea Cavalli (Italian Institute of Technology, Genova, Italy), the whole association process could be simulated. The obtained results indicated that the prebinding of the molecules to the hydrophobic patch on the surface of

hCAII, next to the active site, is stabilized by hydrophobic interactions of the alkyl tails of the ligand molecules. In a subsequent step the ligand diffuses in the final binding position. As a proof of concept the binding of a similar ligand molecule to hCAII which contained a *para* hydroxyethyl chain was simulated by MD experiments as well as measured by SPR. Indeed for this compound the association process was significantly slowed down. Thus, our results are in line with the assumed multi-step process with an intermediate hydrophobic stabilization. Knowledge about the association processes of ligands and proteins is still limited which makes our results a valuable contribution to this field. Especially hCAII is an excellent model system to investigate in the future further inhibitor series by SPR, MD and ITC experiments. ITC experiments should be performed for the here investigated inhibitors to correlate the thermodynamic data with the kinetic properties. This would make this ligand series to one of the first which is characterized by all the above-mentioned techniques.

# 8 Zusammenfassung und Ausblick

Das Ziel der Wirkstoffentwicklung ist das Design von passenden Liganden für das jeweilige Zielprotein. Da die Entwicklung eines neuen Medikamentes bis zu mehr als einer Milliarde Dollar kosten kann, ist die Entscheidung, welche der Kandidaten, die aus einem Screening hervorgegangen sind, weiter verfolgt werden sollen, ein kritischer Schritt. Damit das richtige Molekül ausgewählt werden kann, wird Wissen über den komplexen Erkennungsprozess zwischen Ligand und Protein benötigt. Die thermodynamischen Parameter, besonders der enthalpische Beitrag, zeigen bereits, ohne dass die Komplexkristallstruktur verfügbar sein muss, welche der Inhibitoren die stärksten Interaktionen mit der Proteinbindetasche eingehen. Die kinetischen Daten geben Auskunft über das Assoziations- und Dissoziationsverhalten von Protein-Ligand-Komplexen. Abhängig von der beabsichtigten Indikation kann z.B. ein Ligand mit einer Verweildauer in der Bindetasche, die vorteilhaft für den therapeutischen Erfolg erscheint, ausgewählt werden. In dieser Arbeit wurde die Bindung von Liganden-Serien zu ihrem jeweiligen Zielprotein, ALR2 oder hCAII, mit Röntgenkristallographie, ITC- und SPR-Messungen untersucht, mit dem Ziel die grundlegenden Prinzipen zu verstehen, die auch für ähnliche Wirkstoffdesign-Projekte genutzt werden können.

Für ALR2 konnten mehrere Projekte erfolgreich abgeschlossen werden. ALR2 ist eine interessante Zielstruktur für die Prävention von diabetischen Spätschäden. Jedoch, mit Ausnahme von Epalrestat, welches in Japan und Indien auf dem Markt ist, waren die großen Anstrengung, um einen effizienten und sicheren Hemmstoff für dieses Protein zu finden, bisher nicht erfolgreich. Diese Arbeit trägt mit der Untersuchung des Inhibitors **13**, einem neuen Typ von ALR2-Hemmstoff, zu dieser Suche bei. Das Molekül ist aus einer Datenbank-Suche und nachfolgender Affinitäts- und Selektivitätsbestimmungen, die von Drs. Milan Stefek, Marta Soltesova Prnova und Magdalena Majekova (Bratislava, Slowakei) durchgeführt wurden, hervorgegangen. Dieser Hemmstoff ist selektiv für ALR2 im Vergleich zu zwei verwandten Proteinen, Aldehyd Reduktase und AKR1B10, was notwendig, ist um unerwünschte Nebenwirkungen zu meiden. In dieser Arbeit konnte die Komplexkristallstruktur von **13** mit ALR2 erhalten werden (1,26 Å), welche zu zwei überraschenden Resultaten führte. Zunächst

wurde die Spezifitätstasche der ALR2 in der geschlossenen Form vorgefunden. Dies ist erstaunlich, da gemeinhin die Adressierung dieser Tasche als Voraussetzung für das Design eines selektiven ALR2-Inhibitors gilt. Außerdem wurde ein weiteres Molekül des Liganden am Eingang der Bindetasche, in Wasserstoffbrückenbindungsabstand zum ersten Molekül, gefunden. Die Frage, ob dieses zweite Molekül immer zusammen mit dem ersten bindet, sollte beantwortet werden. Deshalb wurden ITC-Experimente durchgeführt. Diese zeigten, dass **13** mit der gleichen Stöchiometrie bindet, die auch Liganden besitzen, in deren Komplexkristallstrukturen nur jeweils ein Molekül des Liganden aufgefunden wurde. Desweiteren weist die verringerte Besetzung von 70% des zweiten Molekül des Liganden darauf hin, dass das zweite Molekül nur aufgrund der hohen Konzentrationen, die beim Soaking-Prozess genutzt werden, bindet. Zusammengefasst kann **13** aufgrund seiner hohen Affinität und Selektivität als das erste Mitglied einer neuen Gruppe von ALR2-Inhibitoren angesehen werden. Die Affinität dieser Inhibitor-Klasse könnte vermutlich noch weiter erhöht werden, wenn die Spezifitätstasche adressiert werden würde. Weitere Liganden sollten synthetisiert werden und die Komplexkristallstrukturen der Inhibitoren mit den höchsten Affinitäten bestimmt werden.

In einer Kooperation mit den Gruppen von Prof. Alberto Podjarny (CNRS, Illkirch, France), Dr. Ángel R. de Lera (Universidade de Vigo), Drs. Maria Jesús Martin und Jesús Ángel de la Fuente (Biomar Microbial Technologies S. A., León), Drs. Xavier Parés und Jaume Farrés (Universitat Autònoma de Barcelona) wurden zwei Verbindungen (**23** und **24**) untersucht, die ALR2 und das verwandte Protein AKR1B10 inhibieren. Zu diesem Projekt wurde die Messung der thermodynamischen Profile beigesteuert. Im Zuge der Bindung an ALR2 wurde eine Protonenaufnahme von 0.6 Mol Protonen je Mol gebildeten Komplexes ermittelt. Dies stimmt mit bisherigen Ergebnissen an anderen Komplexen vergleichbaren Typs überein. Überraschend wurde für die Bindung an AKR1B10 nur eine Protonenaufnahme von 0.1 Mol Protonen je Mol gebildeten Komplexes gefunden. Dies zeigt, dass sich die Protonierungszustände in der Bindetasche von AKR1B10 vor der Bindung eines Liganden von ALR2 unterscheiden. In einem folgenden Projekt sollte dies näher untersucht werden, z. B. mittels pKa-Wertberechnungen für die Aminosäuren in der Bindetasche von AKR1B10.

Desweiteren wurde eine ALR2-Inhibitor-Serie mit einer Carboxylat-Gruppe und zwei oder drei aromatischen Systemen unter Zuhilfenahme von Röntgenkristallographie, $IC_{50}$-Bestimmungen und ITC-Messungen untersucht. Diese Gruppe stammt ursprünglich von einer virtuellen Leitstruktursuche. In dieser Arbeit wurden die Verbindungen, die aus einem weiteren Optimierungszyklus hervorgingen, untersucht. Die Liganden wurden von der Gruppe von Prof. Martin Schlitzer (Philipps Universität Marburg) synthetisiert. Zu Beginn dieses Projektes wurde die Co-Kristallisation bzw. das Soaking in Citrat-Puffer durchgeführt. Allerdings konnte keiner der untersuchten Liganden unter dieser Bedingung in der Differenzelektronendichte gefunden werden. So musste zunächst ein neues Soaking-Protokoll in TRIS-Puffer entwickelt werden, um die Löslichkeit der Verbindungen zu erhöhen. Dadurch waren Kristallstrukturen im Komplex mit zumindest einigen der Liganden zugänglich. Es konnte gezeigt werden, dass alle Inhibitoren, deren Protein-Ligand-Komplex erhalten wurde, einen sehr ähnlichen Bindemodus aufweisen. Unterschiede konnten in der Position der Carboxylat-Gruppe in der Anionen-Bindetasche gefunden werden. Die meisten Liganden bilden eine H-Brücke mit Trp111NE1 während zwei Inhibitoren ein anderes Interaktionsmuster eingehen. Die letztgenannten Liganden interagieren mit einem Wassermolekül, wodurch die verlorene H-Brücke zu Trp111 ausgeglichen wird. Der erste Protein-Ligand-Komplex mit einem Liganden aus dieser Serie, der ein drittes aromatisches Ringsystem enthält, konnte kristallographisch aufgeklärt werden. Dieser zeigt dass der dritte aromatische Ring nicht die Bindung in einer ähnlichen Orientierung wie für die anderen Mitglieder dieser Gruppe verhindert. Außerdem konnte der erste Protein-Ligand-Komplex eines Inhibitors aus dieser Gruppe mit einer Acrylsäure-Kopfgruppe aufgeklärt werden. Diese führte zu der überraschenden Erkenntnis, dass sich die Konfiguration der in den Liganden vorliegenden Doppelbindung während des Soaking-Prozess von *E* zu *Z* ändert. Offensichtlich kann der Ligand mit der Acrylsäuregruppe in *Z*-Konfiguration vorteilhaftere Wechselwirkungen in der Anionen-Bindetasche eingehen als in *E*-Konfiguration.

In einem weiteren Projekt wurde die Spezifitätstasche der ALR2 mit Hilfe von Derivaten des zuvor beschriebenen Inhibitor IDD393, die in der Gruppe von Prof. Wibke Diederich (Philipps-Universität Marburg) synthetisiert wurden, untersucht. Der Teil dieser Liganden, der die Spezifitätstasche adressiert, wurde variiert. So wurden Inhibitoren erhalten, durch welche

unterschiedliche Konformationen der Spezifitätstasche induziert wurden. Der Vergleich von zwei dieser Liganden, wovon einer die Spezifitätstasche mit einem nicht-substituierten Phenylring adressiert (**49**) und der andere keine Seitenkette enthält, die diese Tasche öffnen konnte (**47**), zeigte, dass die Öffnung der Tasche von einem kleinen aber doch messbaren Verlust in der Enthalpie einhergeht. Da die vergleichende Analyse der jeweiligen hochaufgelösten Proteinkristallstrukturen keine Erklärung für dieses Resultat liefert, kann vermutet werden, dass die Reste der Spezifitätstasche in der offenen Konformation eine höhere Energie besitzen als in der geschlossenen Konformation. Folglich müsste Energie investiert werden, um die Tasche zu öffnen, was den Verlust in der Enthalpie erklären würde. Die Liganden wurden mit verschiedenen Resten versehen, um die Tasche zu adressieren. Eine Nitro-Gruppe, die in *meta* Position an einen Phenylring gebunden ist, führt im Vergleich zum nicht-substituierten Phenylring zu einem starken enthalpischen Gewinn von -35.7 kJ/mol. Zwar wurde bereits zuvor gezeigt, dass ein solcher Substituent in dieser Portion zu einem starken enthalpischen Signal führt, jedoch konnte dieses bisher nicht quantifiziert werden. Der Austausch dieser Nitro-Gruppe durch eine Carboxylat-Funktion führt überraschenderweise zu einem geänderten Bindemodus des Liganden. Der terminale Phenylring der Benzoesäure-Gruppe des Liganden ist nicht in der Spezifitätstasche vergraben, wie es für das Nitrophenyl-Derivat der Fall ist, sondern dieser orientiert sich in Richtung des Lösungsmittel zugänglichen Eingangs der Bindetasche. Um diesen Bindemodus einzunehmen, muss eine interne H-Brücke zwischen dem Amid-Stickstoff und einem der Sauerstoffatome der Carboxylat-Funktion des Liganden, die sonst bei allen anderen Inhibitoren der hier untersuchten Serie vorliegt, aufgegeben werden. Dieser geänderte Bindemodus führt zu einer 2000-fach reduzierten Affinität im Vergleich zum Nitro-Derivat. Vermutlich ist die Desolvation einer Carboxylat-Gruppe deutlich unvorteilhafter als für eine Nitrogruppe, was dieses überraschende Resultat erklären würde. In einem folgenden Projekt sollten weitere Derivate aus dieser Serie untersucht werden, um besser zu verstehen, was den diskriminierenden Faktor für die Öffnung der Tasche darstellt. Dieses hier erhaltende Ergebnis kann auf ähnliche Wirkstoffdesign-Projekte, bei denen die Öffnung einer transienten Subtasche eine Rolle spielt, übertragen werden.

# Zusammenfassung und Ausblick

Die Bindung einer Serie von Benzensulfonamiden an die hCAII wurde mithilfe von SPR-Experimenten und Proteinkristallographie untersucht. Die erhaltenen Resultate wurden mit den Erkenntnissen von MD-Simulationen und zuvor publizierten SFF-Messungen verglichen. Vergleichbar zu den Resultaten der SFF-Messungen konnte gezeigt werden, dass die Assoziationsrate mit der Länge der hydrophoben Alkylkette, die sich in *para*-Stellung zur Sulfonamid-Gruppe befindet, ansteigt. Dies ist zunächst ein überraschendes Resultat, denn zumindest in dem Fall eines rein diffusions-kontrollierten Assoziationsprozesses sollte die Bindung verlangsamt werden, wenn das Molekül eine sterisch anspruchsvolle Seitenkette besitzt. Folglich scheint der Assoziations-Prozess in der hier betrachteten Situation mehrere Schritte zu durchlaufen, von denen einer von einer hydrophoben Interaktion vorteilhaft beeinflusst wird. Die Bestimmung der Kristallstrukturen im Komplex mit dem jeweiligen Inhibitor zeigte, dass alle Liganden einen sehr ähnlichen Bindemodus aufweisen. Allerdings repräsentieren die Kristallstrukturen nur den finalen Bindemodus. Mit MD-Simulationen, die in der Gruppe von Prof. Andrea Cavalli (Italian Institute of Technology, Genua, Italien) durchgeführt wurden, konnte der gesamte Bindeprozess untersucht werden. Dies führte zu der Erkenntnis, dass die zwischenzeitliche Bindung der Moleküle an der hydrophoben Region, die sich in der Nähe der Bindetasche befindet, durch hydrophobe Interaktionen der Alkylkette der Liganden stabilisiert wird. Im nachfolgenden Schritt wandert das Molekül dann in die finale Bindeposition. Um diese Theorie zu überprüfen, wurde die Bindung eines Derivates, das eine Hydroxyethyl-Seitenkette besitzt, mit MD-Simulationen und SPR-Experimenten untersucht. Tatsächlich konnte mit beiden Verfahren gezeigt werden, dass die Bindung dieses Liganden mit der polaren OH-Gruppe an die hCAII stark verlangsamt erfolgt. Folglich deuten unsere Resultate darauf hin, dass die Bindung von Benzensulfonamiden an die hCAII aus mehreren Einzelschritten besteht und eine zwischenzeitliche hydrophobe Stabilisierung involviert ist. Da das Hintergrundwissen über die Assoziation von Proteinen und Liganden noch sehr begrenzt ist, stellt das hier erhaltene Wissen einen wertvollen Beitrag zu diesem Gebiet dar. hCAII kann dabei in zukünftigen Studien als Modellsystem genutzt werden, um weitere Inhibitor-Serien mit SPR-, ITC- und MD-Experimenten zu untersuchen. Vor allem sollten ITC-Messungen an der hier beschriebenen Benzensulfonamid-Serie durchgeführt werden. Dies würde diese Protein-Liganden-Serie zu einer der ersten machen, die mit all den zuvor genannten Techniken untersucht wurde.

# 9 Materials and Methods

## 9.1 Chemicals, Devices, Buffers and Solutions

### 9.1.1 Chemicals

**Table 9-1 Chemicals in alphabetical order.**

| Name | Company | Abbrevation |
|---|---|---|
| 1,4-Dithiothreitol | Roth | DTT |
| 2-(*N*-morpholino)ethanesulfonic acid | Roth | MES |
| 4-(2-Hydroxyethyl)-1-piperazine ethan sulfonic acid | Roth | HEPES |
| Acetic acid 100% | Roth | |
| Activated carbon | Roth | |
| Agar-Agar | Roth | |
| Ammoniumpersulfate | Roth | APS |
| Ammoniumsulfate | Grüssing | |
| Ampicilline | Roth | |
| Bromphenol blue | Merck | |
| Calcium chloride | Roth | |
| Chloramphenicole | Roth | |
| Coomasie Brilliant Blue R250 | Roth | |
| D-(+)- Xylose | Sigma Aldrich | |
| di-Ammonium hydrogen citrate | Sigma Aldrich | |
| Dimethylsulfoxid | Roth | DMSO |
| di-Potassium hydrogen phosphate | Merck | |
| Ethanol 99.8% | Roth | |
| Ethylene diamine tetra acetate | Merck | EDTA |
| Glycerol 99% | Acros Organics | |
| Guanidine-HCL | Merck | |
| Hydrochloric acid 32% | Fisher Chemical | |
| Imidazole | Roth | |
| Isopropanol 99.5% | Acros Organics | |
| Isopropyl-β-D-thiogalactopyranosid | Roth | IPTG |
| Lysozyme | Roth | |
| Methanol | Fisher Chemical | |
| N-(2-Hydroxy-1,1-bis(hydroxymethyl)ethyl)glycine | Merck | Tricine |
| Nickel(II) sulfate | Sigma Aldrich | |

**Continuation of Table 9-1.**

| Name | Company | Abbrevation |
|---|---|---|
| Nicotinamide adenine dinucleotide phosphate | Sigma Aldrich | $NADP^+$ |
| Nicotinamide adenine dinucleotide phosphate, reduced | Roth | NADPH |
| *p*-Chloromercurybenzoic acid | Sigma Aldrich | |
| Peptone from Casein | Roth | |
| Polyethylene glycole 6000 | Sigma Aldrich | PEG 6000 |
| Polyoxyethylene (20) sorbitan monolaurate (Tween 20) | GE Healthcare | |
| Potassium di-hydrogen phosphate | Roth | |
| Rotiphorese® 10x SDS | Roth | |
| Rotiphorese® Gel 30 ( 30% Acrylamide/0.8% Bisacrylamide ) | Roth | |
| Sodium chloride | Roth | |
| Sodium dodecyl sulfate | Sigma Aldrich | SDS |
| Sodium hydroxide | Grüssing | |
| Tris(hydroxamethyl)aminomethane | Roth | TRIS |
| Yeast extract | Roth | |

### 9.1.2 Laboratory devices

**Table 9-2 The devices which were used in this work are listed in alphabetical order.**

| Device | Company |
|---|---|
| 96 well microplates | Greiner Bio-One |
| ÄKTA FPLC | Amersham Biosciences |
| ÄKTA Prime Plus FPLC | GE Healthcare |
| Autoclave T-Line | Fedegari Autoclave Spa |
| Benzamidine FF column | GE Healthcare |
| Biacore S51 SPR instrument | GE Healthcare |
| Biofuge fresco centrifuge | Hereus |
| Centrifuge Avanti® J-25 | Beckmann |
| Centrifuge Biofuge Pico™ | Hereus |
| Centrifuge Multifuge™ 3 | Hereus |
| Glutathione sepharose high performance material | GE Healthcare |
| HERAsafe™ biological safety cabinet | Thermo Scientific |
| HiLoad 26/600 Superdex 200 pg size exclusion column | Amersham Biosciences |
| HiTrap Chelating HP column | GE Healthcare |
| HiTrap DEAE FastFlow sepharose column | GE Healthcare |
| iEMS Reader | MF Labsystems Oy |

**Continuation of Table 9-2.**

| Device | Company |
|---|---|
| IKA® mini G centrifuge | IKA |
| ITC200 microcalorimeter | GE Healthcare |
| Microscope SZ-PT | Olympus |
| Nanodrop™ 2000c | Thermo Scientific |
| PD-10 Desalting Column | GE Healthcare |
| Peristaltic pump P-1 | GE Healthcare |
| pH-meter Five Easy™ FE20 | Mettler Toledo |
| Pipettes | Eppendorf |
| SDS-Page Mini-PROTEAN® Tetra Electrophoresis System | BioRad |
| Shaking incubator Innova 4200 | New Brunswick Scientific |
| Shaking incubator Innova 4300 | New Brunswick Scientific |
| Thermomixer Comfort | Eppendorf |
| Ultrasonic bath Elmasonic S 10 | Elma |
| Ultrasonic Sonifier® 250 | Branson |
| Vivaspin® 10,000 concentrator | Sartorius |

### 9.1.3 Solutions and Buffer

**Table 9-3 Solutions and buffers which were used in this work are listed in alphabetical order.**

| Name | Substances |
|---|---|
| Citrate buffer | 50 mM di-ammonium hydrogen citrate, pH 5.0 |
| Citrate cryo buffer | 120 mM di-ammonium hydrogen citrate, 40% (w/v) PEG 6000, pH 5.0 |
| Citrate stabilization buffer | 120 mM di-ammonium hydrogen citrate, 25% (w/v) PEG 6000, pH 5.0 |
| Coomasie blue destaining solution | 10% (v/v) acetic acid, 40% (v/v) methanol |
| Coomasie blue staining solution | 10% (v/v) acetic acid, 40% (v/v) methanol, 1 g/L Coomasie Brilliant Blue R250 |
| Elution buffer (ALR2) | 200 mM imidazole, 16 mM TRIS and 400 mM NaCl, pH 8.0 |
| hCAII well solution | 2.5- 3.0 M $(NH_4)_2SO_4$, 100 mM TRIS, saturated with *p*-chloromercurybenzoic acid, pH 7.8 |
| ITC buffer (ALR2) | 10 mM HEPES, Tricine or TRIS, pH 8.0 |
| LB media | 1% (w/v) NaCl, 1% (w/v) pepton, 0.5% (w/v) yeast, autoclave |
| MES cryo buffer | 50 mM MES, 40% (w/v) PEG 6000, pH 5.5 |
| MES stabilization buffer | 50 mM MES, 25% (w/v) PEG 6000, pH 5.5 |
| PBS buffer | 139 mM NaCl, 2.7 mM KCl, 10.0 mM $Na_2HPO_4$, 1.8 mM $KH_2PO_4$, pH 7.4 |

**Continuation of Table 9-3.**

| Name | Substances |
|---|---|
| SDS loading gel buffer | 1.0 M TRIS, pH 6.8 |
| SDS sample buffer 4x | 250 mM TRIS, 200 mM DTT, 0.04% (w/v) bromphenol blue, 8% (w/v) SDS, 40% (v/v) glycerol, pH 6.8 |
| SDS separation gel buffer | 1.0 M TRIS, pH 8.8 |
| Sodium phosphate buffer | 18.4 mM di-potassium hydrogen phosphate, 25.3 mM potassium dihydrogen phosphate, pH 6.2 |
| Super LB media | 0.5% (w/v) NaCl, 2.5% (w/v) pepton, 1.5% (w/v) yeast, autoclave |
| TRIS 1 | 20 mM TRIS and 500 mM NaCl, pH 8.0 |
| TRIS 2 | 10 mM TRIS, pH 8.0 |
| TRIS cryo buffer | 100 mM TRIS, 40% (w/v) PEG 6000, pH 8.0 |
| TRIS stabilization buffer | 100 mM TRIS, 25% (w/v) PEG 6000, pH 8.0 |
| Washing buffer (ALR2) | 20 mM TRIS, 10 mM imidazole, 500 mM NaCl, pH 8.0 |

## 9.2 Experimental Procedures Regarding ALR2

### 9.2.1 Expression and purification

The expression and purification followed previously described protocols.[134, 144, 199] A plasmid (pET15b, Novagen) which contained the open reading frame for the ALR2 gene together with a hexahistidine tag was kindly provided by Alberto Podjarny (IGBMC, Illkirch, France). The used ALR2 gene was sequenced. The result is shown in Figure 10-9 in the Appendix. The fusion protein was expressed in *E. coli* strain BL21 gold (DE3). An overnight culture of LB medium (100 mL) with 0.1 mg/mL ampicillin was inoculated and incubated for 16 hours at 310 K and 200 rpm. 25 mL of the suspension was added to 1 L of Super-LB-medium with 0.1 mg/mL of ampicillin. A total volume of 4 L media has been used. The cells were grown to an $OD_{600}$ of around 3.0 at 310 K and 200 rpm. Afterwards 1 mL of 1 M IPTG was added to 1 L of cell culture. The protein was expressed for 3 hours at 310 K and 200 rpm. The cells were then harvested by centrifugation with 6,000 rpm at 277 K for 10 minutes. The supernatant was discarded. The pellet of 1 L of Super LB medium was resuspended in a total volume of 25 mL of TRIS 1 (20 mM TRIS, 500 mM NaCl, pH 8.0). The cell walls were disrupted using ultrasound (six times for 1 min, duty cycle: 70%, micro tip limit: 3). During this procedure the suspension was kept on ice. Afterwards the suspension was centrifuged for 45 minutes at 20,000 rpm. The supernatants of all tubes were combined. A HiTrap Chelating HP column (GE Healthcare) was connected to an Äkta FPLC system (Amersham Pharmacia Biosciences). The protein solution was loaded on the column with a flow rate of 2.5 mL/min. TRIS 1 was injected until low absorption values were reached. Afterwards a short washing step (30 s, 2 mL/min) with washing buffer (20 mM TRIS, 10 mM imidazole, 500 mM NaCl, pH 8.0) was performed. The fusion protein was eluted using a linear gradient of elution buffer (200 mM imidazole, 16 mM TRIS, 400 mM NaCl, pH 8.0). Using a Vivaspin 10,000 ultracentrifugation device (Sartorius) the buffer was exchanged to TRIS 2 (10 mM TRIS, pH 8.0) and concentrated to 15 mL. $CaCl_2$ was added to a final concentration of 5 mM. For the cleavage of the hexahistidine tag thrombin (kindly provided by CSL Behring, Marburg) was added and the solution was stored for at least 20 h at 277 K. The solution was diluted by adding TRIS 2 to a volume of 150 mL. A 5 mL HiTrap DEAE FastFlow sepharose column (GE Healthcare) was connected to an Äkta FPLC. The

column was loaded with the protein solution (2 mL/min) and afterwards washed with TRIS 2. When low absorption values were reached the protein was eluted with a linear gradient of TRIS 1 (1 mL/min).

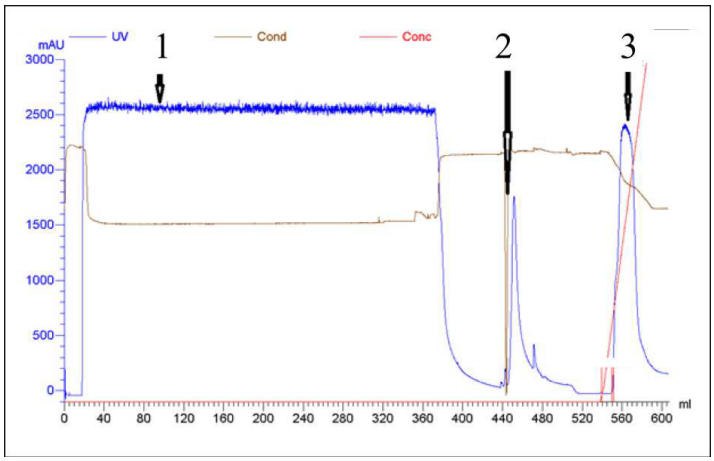



**Figure 9-1 Purification of ALR2.** Plateau 1: Loading of HiTrap Chelating HP column with the supernatant; Peak 2: Short washing step; Peak 3: Elution from the column using an imidazole gradient.

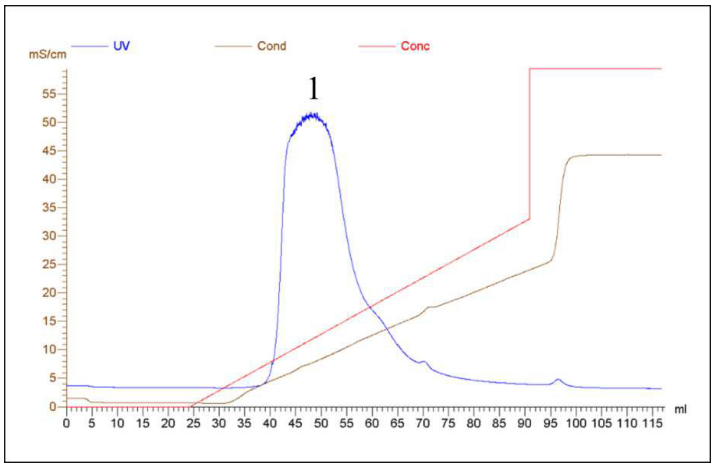



**Figure 9-2 ALR2 eluted from HiTrap DEAE FastFlow sepharose column using a NaCl-gradient (Peak 1).**

The identity and purity of the protein was verified by SDS gel electrophoresis. Finally, the buffer was exchanged using a Vivaspin 10,000 ultracentrifugation device (Sartorius), a PD10 column or dialysis to the needed buffer and concentrated to the required concentration.

.

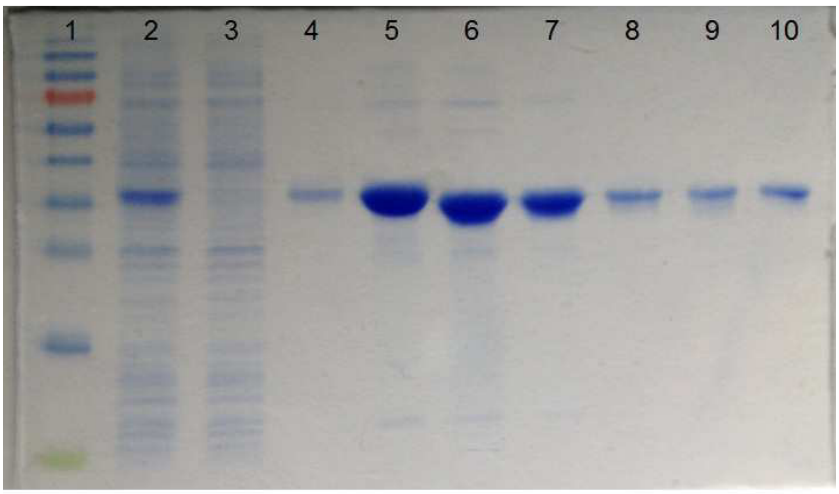



**Figure 9-3 SDS gel of a typical purification of ALR2. 1**:marker; **2**:after lysis and centrifugation; **3**:flowthrough of the HiTrap Chelating HP column; **4**:peak after injection of washing buffer; **5**:protein after HiTrap Chelating HP column; **6**:ALR2 after cleavage; **7**: ALR2 after DEAE column; **8** and **9**: different batches of purfied ALR2; **10**: inhouse reference of ALR2.

### 9.2.2 $IC_{50}$ value determination

The determination of $IC_{50}$ values followed previously described protocols.[144,199,200] After the purification, the protein buffer was exchanged to sodium phosphate buffer (di-potassium hydrogen phosphate, potassium dihydrogen phosphate, pH 6.2) using a PD10 column. The measurement of the $IC_{50}$ value is based on the decrease of the NADPH concentration during the conversion of D-Xylose to Xylitol performed by ALR2. NADPH is converted to the oxidized form $NADP^+$. The level of NADPH was monitored through its absorbance at 340 nm which was measured by a plate reader at 310 K (iEMS Reader, MF, Labsystems Oy, Finnland). $NADP^+$ does not absorb at this wavelength. 96 well microplates (Greiner Bio-One) were used. Each well contained 250 µL of the reaction mixture consisting of 0.28 µM ALR2, 0.14 mM NADPH and 26.5 mM D-Xylose in sodium phosphate buffer at pH 6.2. The concentration of the ALR2 stock was adjusted to 0.125 mg/mL. 20 different concentrations of the inhibitor were measured in parallel. The reaction mixture contained 5% (v/v) DMSO. Only a few investigated inhibitors made the addition of 10% (v/v) DMSO necessary. In each column of the plate two wells were

used to monitor the uninhibited reaction. NADPH and the protein were added in the last step to start the reaction. A multipipette was used to mix the solution in each well because a comparable surface shape of the drops was found to be a prerequisite for stable measurements. The rate of the reaction in each well with an inhibitor was divided through the rate of the uninhibited reaction. This ratio was plotted on the y-axis, while on the x-axis the concentration of the inhibitor was plotted in logarithmic scale. The range of investigated concentrations of the inhibitor was adjusted in a way that the resulting curve had a sigmoidal shape. The program GraFit (Version 4.09, Erithacus Software Limited) was used to calculate the $IC_{50}$ value. Each measurement was performed at least in triplicate. The mean values and the standard deviations were calculated.

### 9.2.3 Crystallization of ALR2

#### 9.2.3.1 Crystallization of *holo*-crystals in citrate buffer

After purification the protein was transferred into 50 mM diammonium hydrogen citrate pH 5.0 using Vivaspin 10,000 ultracentrifugation device or dialysis. The concentration of the protein was adjusted to 30 mg/mL. The crystallization solution was prepared as shown in Table 9-4.

**Table 9-4 Preparation of the ALR2 crystallization solution in citrate buffer.**

| | |
|---|---|
| **Mixture 1** | |
| 50% (w/v) PEG 6000 | 340 µL |
| 1 M DTT | 34 µL |
| 50 mM diammonium hydrogen citrate pH 5.0 | 1335 µL |
| **Mixture 2** | |
| 30 mg/mL ALR2 in 50 mM diammonium hydrogen citrate pH 5.0 | 100 µL |
| 200 mg/mL $NADP^+$ in 50 mM diammonium hydrogen citrate pH 5.0 | 0.66 µL |
| **Crystallization Solution** | |
| Mixture 1 | 100 µL |
| Mixture 2 | 100 µL |

Aliquots of 200 µL of the crystallization solution were flash-frozen using liquid nitrogen. The final concentrations in the crystallization solution were 15 mg/mL ALR2, 5% (w/v) PEG 6000, 5.2 mg/mL DTT and 0.7 mg/mL $NADP^+$. The hanging drop method was used. The well solution

contained 120 mM diammonium hydrogen citrate (pH 5.0) with 20% (w/v) PEG 6000. 9 µL of the crystallization solution were equilibrated for at least 3 days at 277 K. Afterwards streak seeding with a horse hair was performed. The crystallization plate was transferred to 291 K directly after seeding. Crystals appeared after 1-3 days and reached their maximum size after two weeks.

#### 9.2.3.2 Co-crystallization in citrate buffer

For co-crystallization experiments 9 µL of the crystallization solution were mixed with 3 µL of a solution containing 5 mM of the respective inhibitor in 120 mM diammonium hydrogen citrate, 25% (w/v) PEG 6000, 5% DMSO, pH 5.0. The crystal structure ALR2-**48** was obtained using this procedure.

#### 9.2.3.3 Crystallization in MES buffer

Mixture 1 contained 100 µL of 30 mg/mL ALR2 in 50 mM MES at pH 5.5 with 0.75 µL of 200 mg/mL NADPH and 0.1 µL 1 M DTT. Mixture 2 contained 20 mM ammonium sulfate, 1 mM DTT, 50 mM MES and varying concentrations of PEG 6000 (10-20%) at pH 5.5. 10 µL of mixture 1 and 2 were mixed. As well solution 120 mM MES with 20% (w/v) PEG 6000 at pH 5.5 was used. The droplets were equilibrated for 7 days at 295 K. Afterwards streak seeding with *holo*-crystals from the citrate condition was done. The crystals appeared after several days and reached their full size after three weeks.

#### 9.2.3.4 Soaking experiments

New soaking procedures using MES or TRIS buffer were developed in the course of this work. In each case the *holo*-crystals were grown in citrate buffer.

#### 9.2.3.5 Soaking in citrate buffer

A saturated solution of the inhibitor in citrate stabilization buffer (120 mM diammonium hydrogen citrate pH 5.0 with 25% (w/v) PEG 6000) was prepared. The crystals were transferred into the soaking drop and fished after one to three days. The complex crystal structure ALR2-**51** was obtained using this procedure.

### 9.2.3.6 Soaking in MES buffer

To prevent the crystals from damage the transfer into the MES buffer was performed in a stepwise fashion. A saturated solution of the inhibitor in MES stabilization buffer (50 mM MES, pH 5.5 with 25% (w/v) PEG 6000) was prepared and mixed with citrate stabilization buffer as shown in Figure 9-4. The crystals were equilibrated in each drop for several hours. Afterwards the crystals were transferred for one to three days into the final drop which consisted of the saturated solution of the inhibitor in MES stabilization buffer. The complex crystal structures ALR2-**30** and ALR2-**37** were obtained using this procedure.

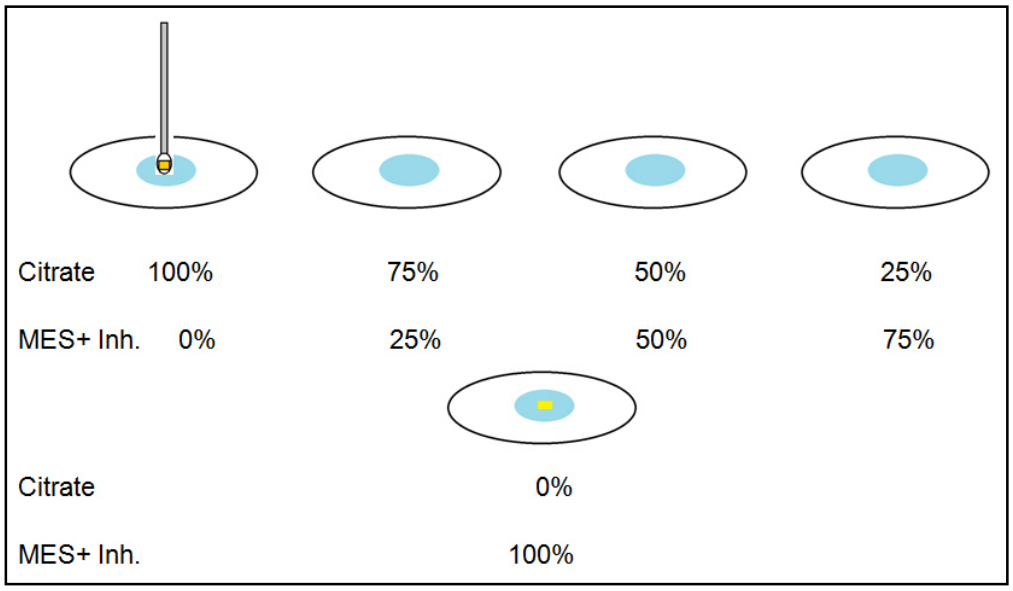



**Figure 9-4 Soaking experiments in MES buffer.** Citrate: citrate stabilization buffer; MES: MES stabilization buffer; Inh.: Inhibitor.

### 9.2.3.7 Soaking in TRIS buffer

The same protocol as described for the soaking experiments in MES buffer was applied. Here, TRIS stabilization buffer (100 mM TRIS, 25% (w/v) PEG 6000, pH 8.0) was used. The complex crystal structures of ALR2 with **13**, **31**, **43**, **47**, **49**, **50**, **52** and **53** were obtained using this procedure.

#### 9.2.3.8 Cryoprotection of the crystals

120 mM di-ammonium hydrogen citrate (pH 5.0), 50 mM MES buffer (pH 5.5) or 100 mM TRIS buffer (pH 8.0), respectively, mixed with 40% (w/v) PEG 6000, were used to cryoprotect the crystals. 1 mM of the corresponding inhibitor with 1% (v/v) DMSO were added. The crystals were briefly dipped into the cryo buffer and were afterwards flash-frozen in liquid nitrogen.

### 9.2.4 Structure determination and refinement

All datasets were collected at the synchrotron BESSY II in Berlin, Germany,[201] at beamline 14.1 or 14.2 (Table 10-3 in the Appendix). The datasets were processed and scaled with HKL2000[137] or XDS.[202] Molecular Replacement was done with Phaser[138] as implemented in the CCP4[203] suite using the PDB entry 2DUX[204] as a model. Using the program Phenix.refine (version 1.8.4_1496) of the Phenix suite[141] three cycles of rigid body refinement, five cycles of simulated annealing and five cycles of refinement were performed. Afterwards alternating cycles of model building in Coot[139] and five cycles of refinement in Phenix were done. The temperature factors of all atoms except hydrogens were anisotropically refined. The 3D coordinates of the inhibitor molecules were built with MOE[205] and afterwards energy minimized using the Tripos force field in Sybyl[206] (Tripos International, version 2.0). The restraints file was generated using eLBOW from the Phenix suite.[142] The final model was validated using MolProbity[207] as implemented Phenix. The final models were deposited in the PDB.[208] Statistics for data collection and refinement are shown in Table 10-3 in the Appendix.

### 9.2.5 ITC measurements

The measurements were performed with an $ITC_{200}$ (GE Healthcare) at 298 K in HEPES buffer. In selected cases the overlaid protonation effect was investigated. In these cases the measurements were done in HEPES, Tricine and TRIS buffer. The cell was filled with a solution of ALR2 in the corresponding buffer with an excess of $NADP^+$ and 3% DMSO. The syringe was filled with a solution containing the same concentration of $NADP^+$ and DMSO in the corresponding buffer and a concentration of the inhibitor which exceeded 10-20 times the concentration of the protein in the cell. The concentrations of the protein and the inhibitor were adjusted to obtain curves with a sigmoidal shape. An initial injection of 0.3-0.5 µL which was omitted during data evaluation was followed by injections of 1.0-2.0 µL with spacings of 180-300 s in between.

For ligands with a too weak or too strong binding affinity the establishment of displacement titrations was necessary (see Section 5.7.2).

For the investigation of the overlaid protonation effect the measurements were done in three different buffer systems (HEPES, Tricine and TRIS). This procedure was previously described in literature.[42, 88] The measured enthalpy ($\Delta H^0_{obs}$) was plotted on the y-axis and the ionization enthalpy of the corresponding buffer system ($\Delta H^0_{ion}$) on the x-axis. The values for $\Delta H^0_{ion}$ were taken from literature.[87, 88] The slope of the regression line indicates how many protons per mole of formed complex are picked up (positive slope) or released (negative slope). The interception point with the y-axis gives the enthalpy corrected for the protonation effect ($\Delta H^0_{bind}$). The resulting plots are shown in Figure 2-6 (Chapter 2), Figure 3-5 (Chapter 3) and Figure 5-10 (Chapter 5).

The peaks of the thermograms were integrated using NITPIC,[160] the isotherms were fitted with SEDPHAT[161] and the curves were plotted using GUSSI.[209] Representative thermograms and regression curves of all investigated inhibitors are shown in Figures 10-1 - 10-4 in the Appendix.

## 9.3 Experimental Procedures Regarding hCAII

### 9.3.1 Introductory remarks

The following description of the methods used in the hCAII project was taken from the same manuscript which was prepared for submission to a scientific journal as chapter 6. Additionally, Figure 9-5 and Figure 9-6 were added. The methods for protein expression, purification, crystallization and soaking followed protocols previously described by Schulze Wischeler and coworkers.[210]

#### 9.3.1.1 Expression and purification of hCAII

100 mL LB media containing 100 µg/mL ampicillin (Roth) and 34 µg/mL chloramphenicol (Roth) were inoculated with BL21 Codon Plus cells containing a pGEX-4T1 (GE Healthcare) plasmid with the sequence of the hCAII-glutathione-S-transferase fusion protein. The used hCAII gene was sequenced. The result is shown in Figure 10-10 in the Appendix. After shaking with 200 rpm for 16 h at 310 K, 25 mL of this overnight culture were transferred into 1L LB media (with 100 µg/mL ampicillin). At an $OD_{600}$ of 0.6-0.8, 1 mL of a 2 M IPTG solution was added. After 4 h at 305 K, the cell suspension was centrifuged at 5,000 rpm. The cells were resuspended in PBS buffer (139 mM NaCl, 2.7 mM KCl, 10.0 mM $Na_2HPO_4$, 1.8 mM $KH_2PO_4$, pH 7.4) and the cell walls disrupted using ultrasound. The suspension was centrifuged at 20,000 rpm. A column which was filled with 20 mL glutathione sepharose high performance material (GSH column, GE Healthcare) was equilibrated with reduced glutathione (Sigma Aldrich) followed by PBS buffer and afterwards loaded with the supernatant. The column was washed with 500 mL PBS buffer. Afterwards a solution of thrombin (CSL Behring, Marburg) was injected onto the column. Overnight, the fusion protein was cleaved on the column. The released hCAII was washed from the column with PBS buffer. To remove the thrombin, a HiTrap 1 mL Benzamidine FF column (GE Healthcare) was used (Figure 9-5). A final purification step with a HiLoad 26/600 Superdex 200 pg size exclusion column (Amersham Biosciences) which was pre-equilibrated with 50 mM TRIS buffer (pH 7.8) was performed (Figure 9-6). The purity of the protein was analyzed by SDS gel electrophoresis.

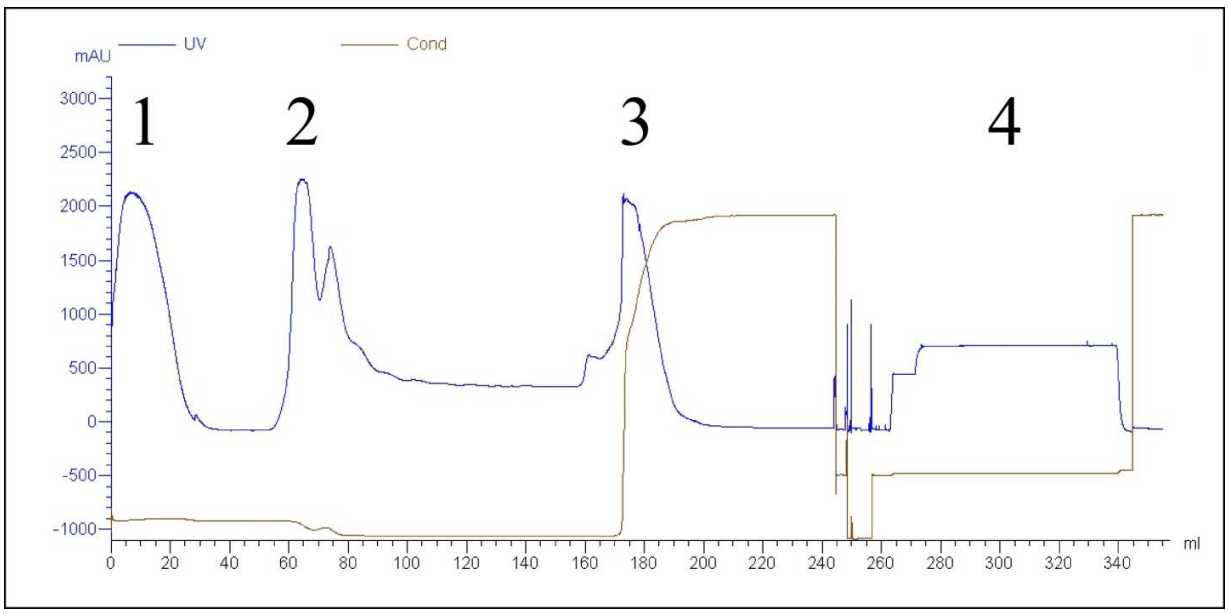



**Figure 9-5 First purification steps of hCAII**. Peak 1: Elution of cleaved hCAII from the GSH column. Peak 2: Washing of GSH column with 10 mM reduced glutathione. Peak 3: Washing of GSH column with 6 M guanidine-HCL. Peak 4: Purification using Benzamidine FF column

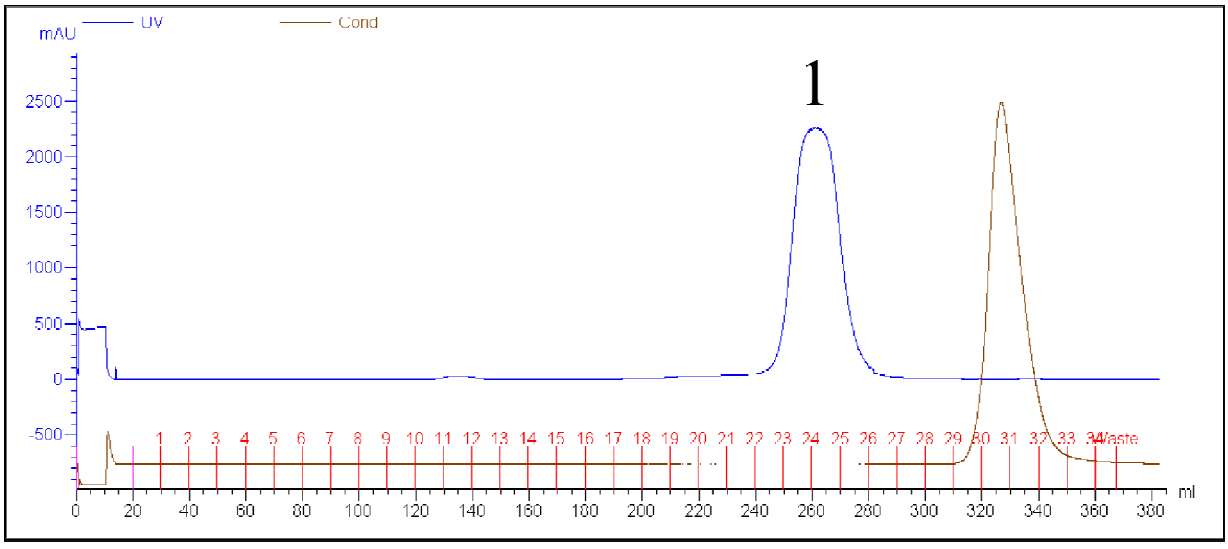



**Figure 9-6 Purification of hCAII using a HiLoad 26/600 Superdex 200 pg size exclusion column (Peak 1).**

#### 9.3.1.2 Crystallization and soaking

The concentration of hCAII in 50 mM TRIS pH 7.8 was adjusted to 10 mg/mL. 2.5 µL of this solution was mixed with 2.5 µL of the well solution (2.7 M $(NH_4)_2SO_4$, 100 mM TRIS, saturated with *p*-chloromercurybenzoic acid) and placed on a cover slide as a hanging drop. Crystals appeared overnight and reached their full size after 2 weeks. The crystals were soaked with a saturated solution of the inhibitor in a stabilization buffer (3.0 M $(NH_4)_2SO_4$, 100 mM TRIS) for one day. Afterwards they were briefly dipped in the cryo buffer (2.7 M $(NH_4)_2SO_4$, 100 mM TRIS, 25% (v/v) glycerol, 1 mM of the respective inhibitor, 1% (v/v) DMSO) for a short time and subsequently frozen in liquid nitrogen.

#### 9.3.1.3 Data collection, processing and refinement

All data sets were collected at beamline 14.1 or 14.2 (Table 10-4 in the Appendix) at the synchrotron BESSY II (Berlin-Adlershof, HZB, Germany).[201] The data sets were processed and scaled using XDS.[202] Molecular replacement was done using Phaser [138] from the CCP4 suite [203] using PDB entry 1CNI as a starting model.[211] 5% of the reflections were marked for the calculation of the $R_{free}$ value. In the Phenix suite[141] the program PHENIX.refine (version 1.8.4-1492) was used to perform three cycles of rigid body refinement and five cycles of Cartesian simulated annealing followed by five cycles of refinement. Then, alternating cycles of model building in Coot[139] and refinement in Phenix were performed. Single atoms of side chains of amino acids were deleted if they were not clearly visible in the $2F_o$-$F_c$ electron density at 1.0 σ. Water molecules were only placed when indicated by the $F_o$-$F_c$ difference electron density at 3.0 σ. The tails of ligands **56**-**58** were not immediately visible. After assigning the core structure of the respective ligand to the density, they became visible atom by atom in following refinement cycles in the $F_o$-$F_c$ electron density maps at 3.0 σ. An anisotropic refinement of the temperature factors for all atoms except hydrogens was performed. The coordinates and restraints for the ligands and for p-chloromercurybenzoic acid were generated using Grade.[212] Finally, riding hydrogen atoms were added to the model using the riding model. However, since the $R_{free}$ value did not improve significantly we refrained from adding hydrogen atoms to the final model. Validation was performed using MolProbity[141] as implemented in the Phenix suite and PROCHECK.[213] The coordinates were deposited in the PDB. Statistics for data collection, processing and refinement are shown in Table 10-4 in the Appendix.

#### 9.3.1.4 SPR measurements

The measurements were performed on a Biacore S51 device (GE Healthcare, Uppsala). For immobilization, PBS buffer (139 mM NaCl, 2.7 mM KCl, 10.0 mM $Na_2HPO_4$, 1.8 mM $KH_2PO_4$) with 0.05% Tween 20, pH 7.4, was used as a background buffer. The standard amine coupling procedure was applied. A series S Sensor Chip CM5 was activated with a mixture of 0.1 M NHS (N-hydroxysuccinimide) and 0.4 M EDC (1-ethyl-3-(3-dimethylaminopropyl) carbodiimide-hydrochloride) at a flow rate of 10 µL/min for 10 minutes. Then, a 50 µg/mL solution of hCAII in 10 mM sodium acetate (pH 5.2) with 30 µM ethoxzolamide was injected for 2 minutes followed by an injection of 1 M ethanolamine hydrochloride (pH 8.5) for 7 minutes. This procedure led to immobilization levels between 3500 and 4900 RU. For measuring the compounds, PBS buffer with 0.05% Tween 20 and 2% DMSO, pH 7.4, was used as the running buffer. The compounds were injected at a flow rate of 30 µL/min for 120 s. Then, the dissociation was monitored for 300 s. No regeneration was needed since all the investigated compounds completely dissociated from the sensor chip surface in the applied time window. In total, 10 concentrations were investigated for each compound in a twofold dilution series. For the different compounds, the following concentrations were measured: **54** (19.5 nM-10.0 µM), **55** (9.8 nM-5.0 µM), **56** (3.9 nM-2.0 µM), **57** (1.2 nM-600 nM), **58** (9.8 nM–5.0 µM), **59** (31.3 nM-16 µM) and **60** (273.4 nM-140.0 µM). A solvent correction cycle using DMSO concentrations between 1.4–2.8% was run in between every two sample concentration series. The data were evaluated with the Biacore S51 evaluation software (version: 1.2.1). The signals of the reference spot were subtracted from the signals of the measurement spots. Then, the sensorgrams were corrected for mismatches in the DMSO concentrations. Two blank injections, carried out contemporary to the injections of the respective ligand, were subtracted from the sensorgrams. For each single measurement, the most robust method for fitting was chosen with a global maximum response ($R_{max}$) and a constant value of 0 for the shift of the refractive index (RI) at the beginning and end of each injection. The kinetic parameters $k_{on}$ and $k_{off}$ were obtained using a simple 1:1 interaction model including a term for a mass transfer limitation as implemented in the evaluation software. The test for a mass transfer limitation as implemented in the evaluation software showed that the curve that was fitted into the sensorgrams was not mass-transfer-limited. Thus, the obtained kinetic constants of our final results were reliable. $K_D$ (kinetic) was

obtained by $k_{off}/k_{on}$. $K_D$ (steady state) was obtained by plotting the binding level at the end of each injection against the concentration of the compound and fitting it to a single-site binding isotherm.

# 10 Appendix

## 10.1 ITC Example Curves and Fitted Regression Curves

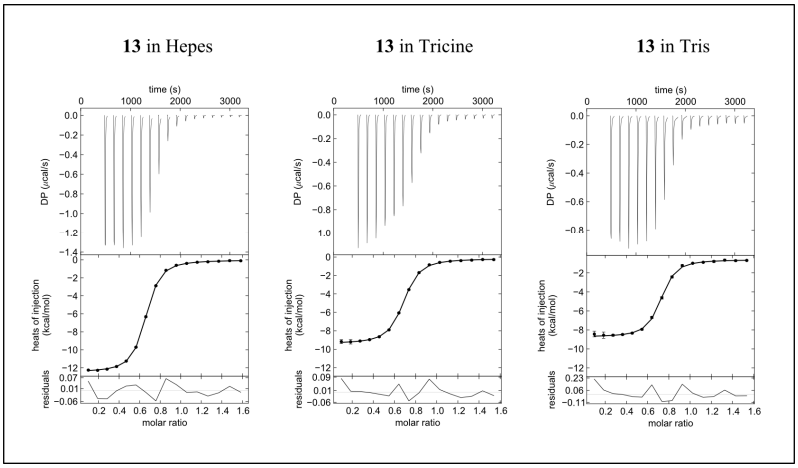



**Figure 10-1 Representative thermograms and fitted regression curves for 13.** The plots were generated using GUSSI.[209]

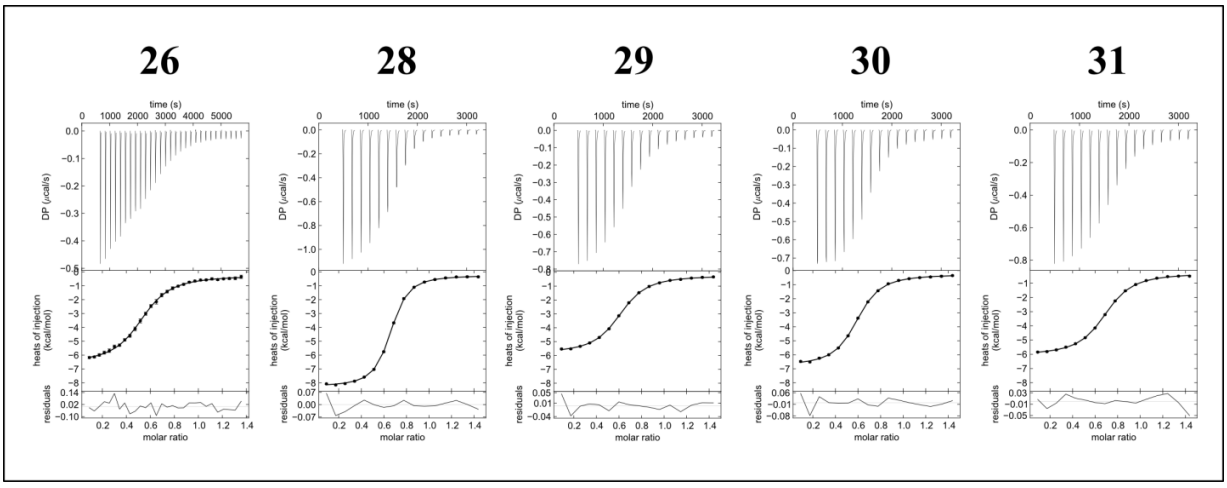



**Figure 10-2 Representative thermograms and fitted regression curves for 26, 28, 29, 30 and 31.** All measurements were performed in HEPES. The plots were generated using GUSSI.[209]

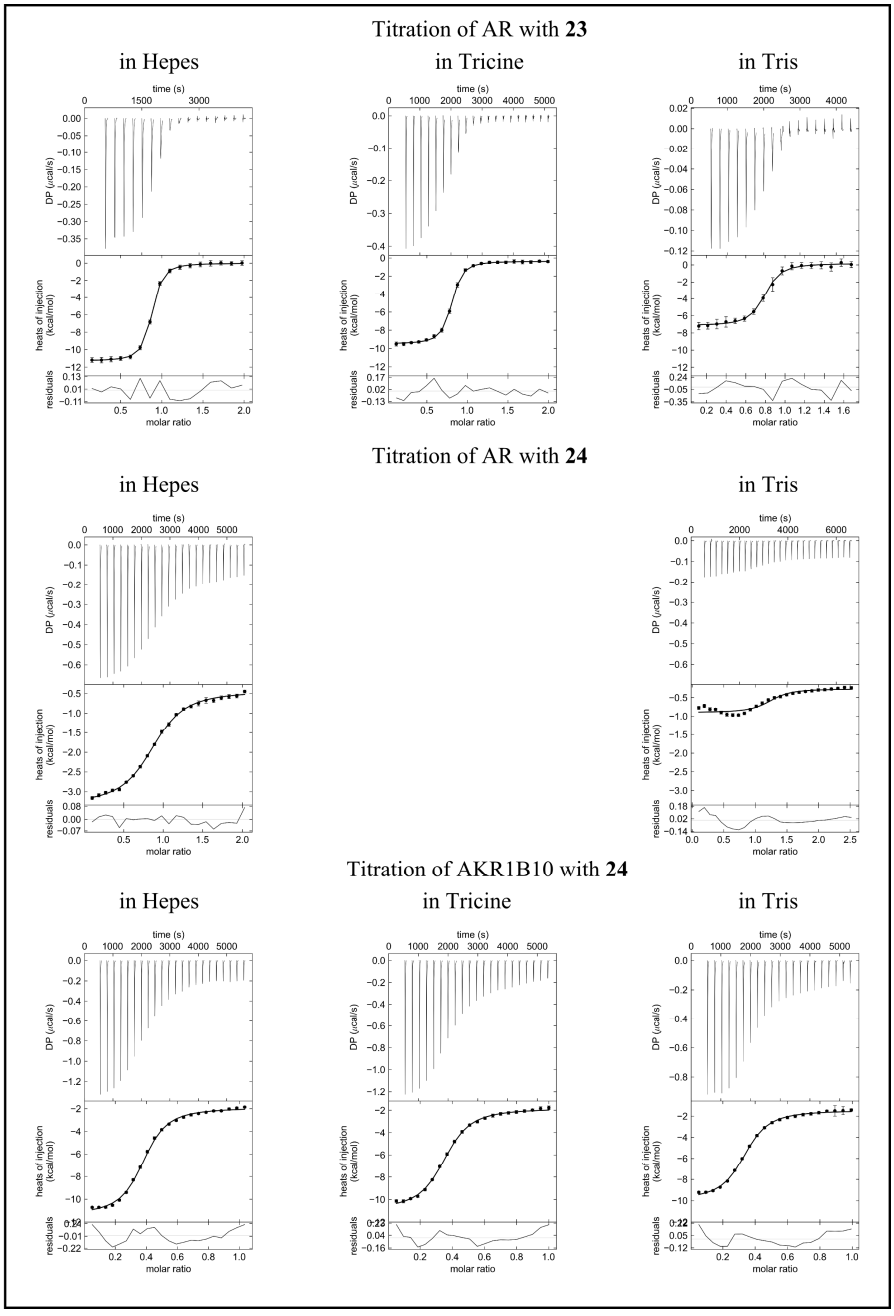



**Figure 10-3 Representative thermograms and fitted regression curves for 23 and 24.** The plots were generated using GUSSI.[209]

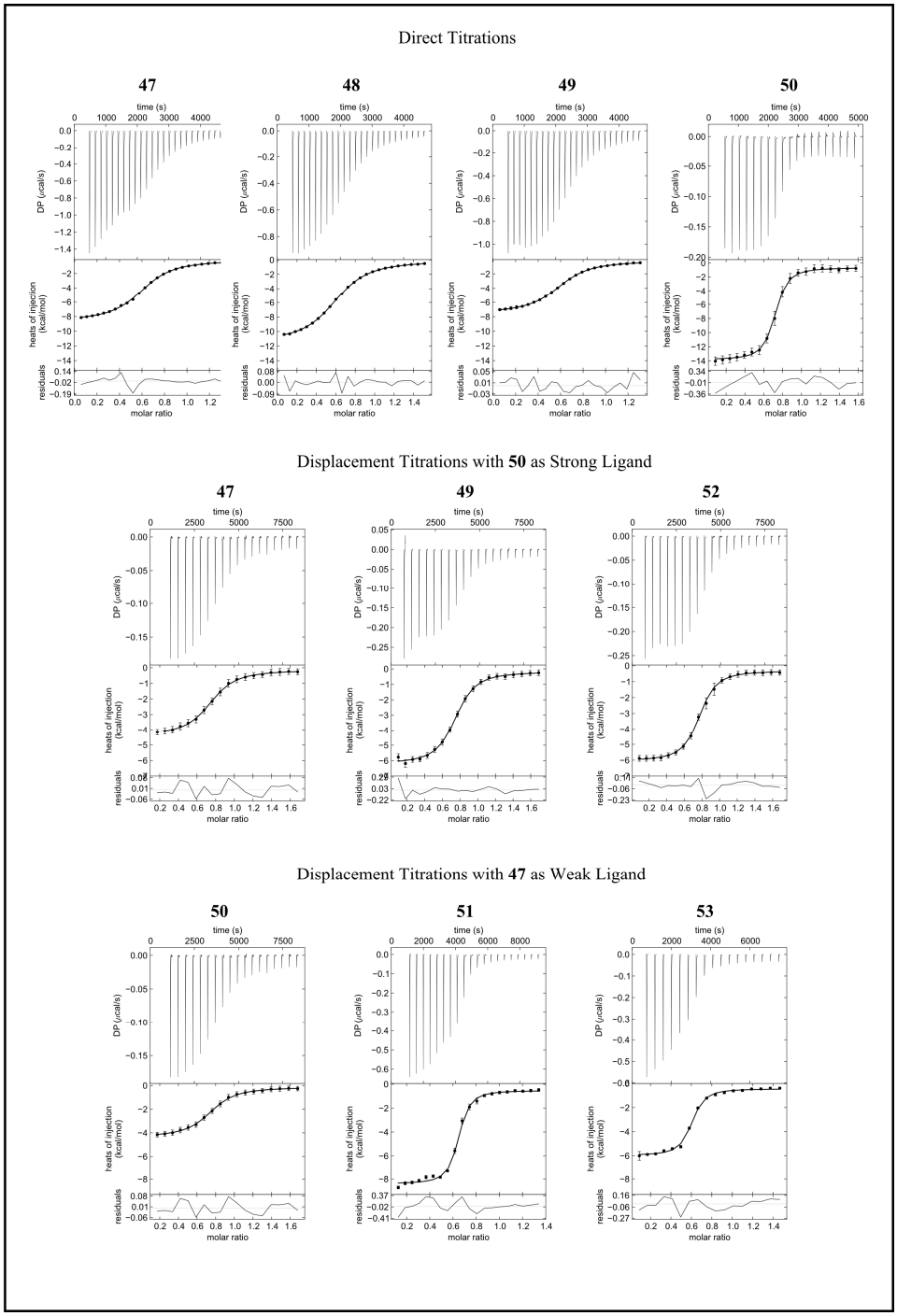



**Figure 10-4 Representative thermograms and fitted regression curves for 47-53.** All measurements were performed in HEPES. The plots were generated using GUSSI.[209]

## 10.2 SPR Sensorgrams

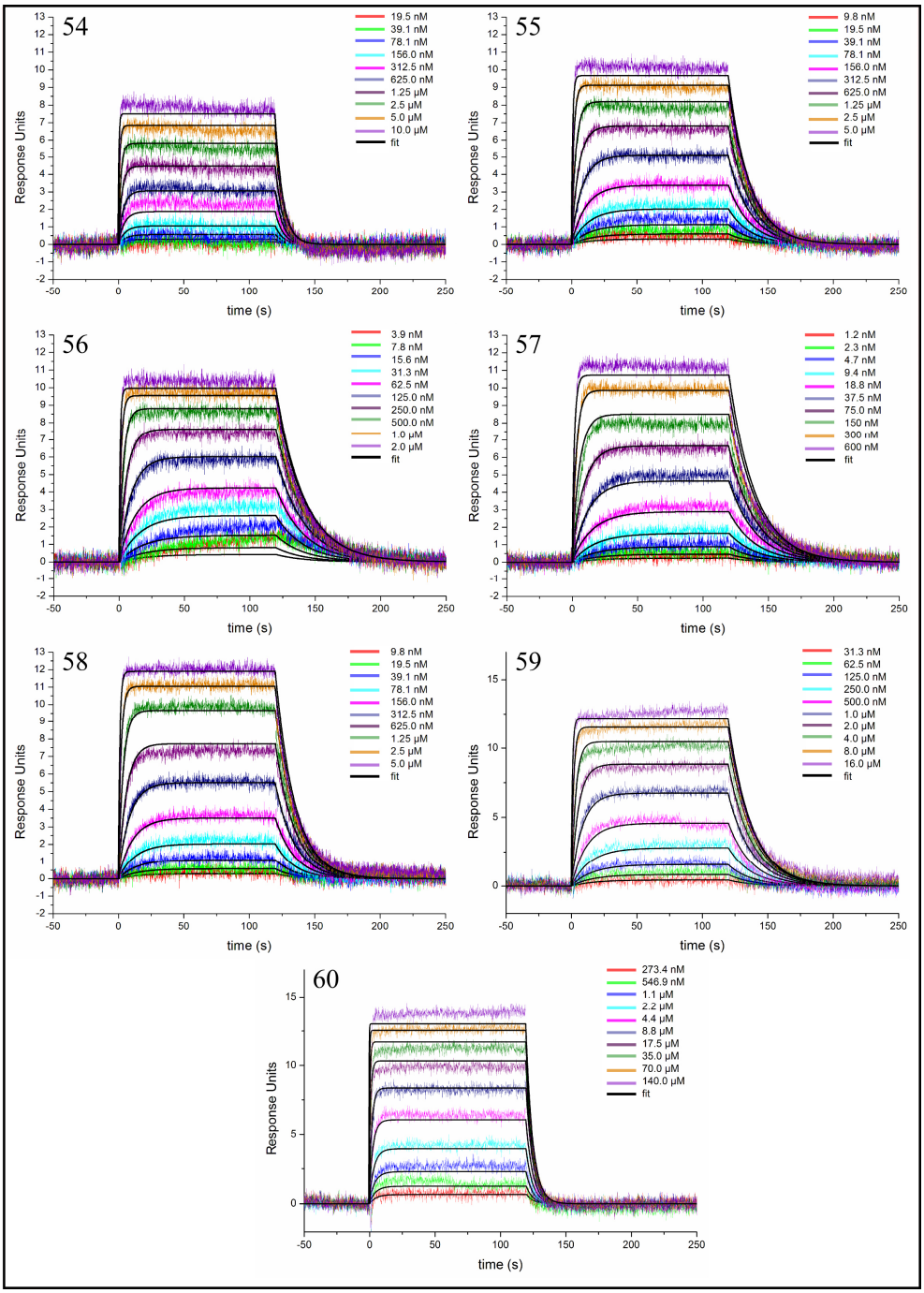



**Figure 10-5 Example sensorgrams for 54-60.** The curves for **54**-**58** were measured with a 3700 RU surface, the curves for **59** and **60** with a 4500 RU surface.

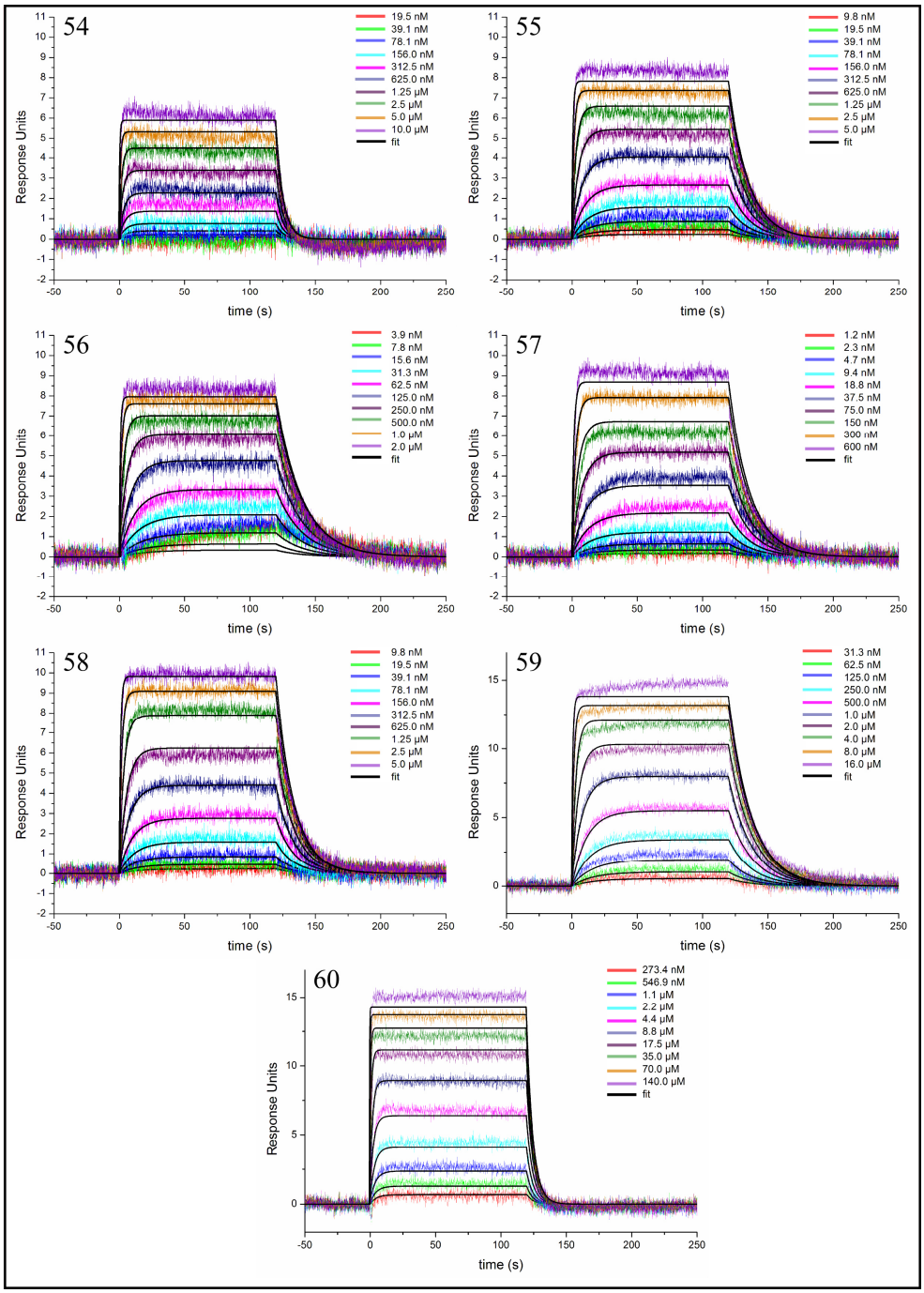



**Figure 10-6 Example sensorgrams for 54-60.** The curves for **54**-**58** were measured with a 3600 RU surface, the curves for **59** and **60** with a 4500 RU surface.

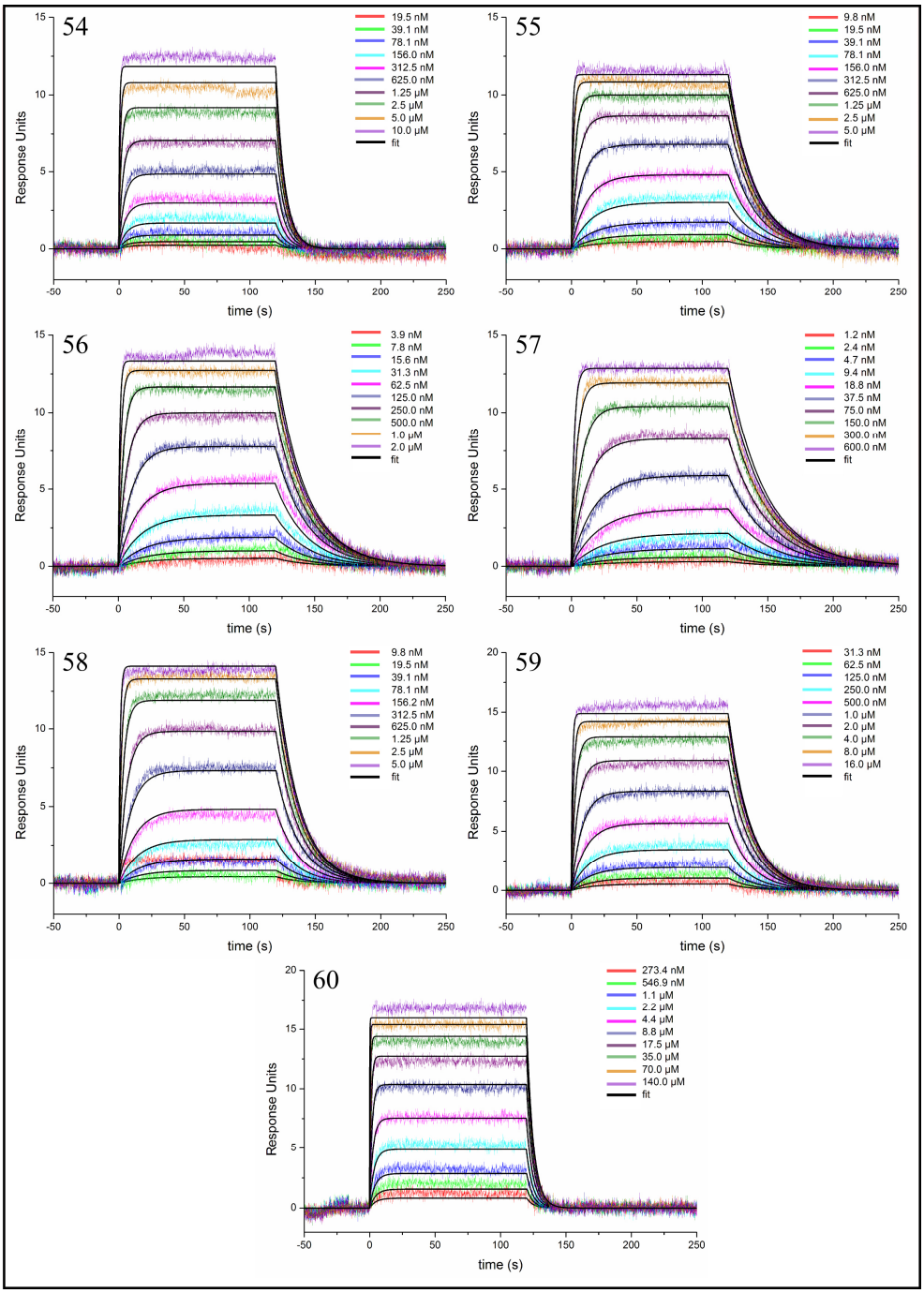



**Figure 10-7 Example sensorgrams for 54-60.**The curves for **54**-**58** were measured with a 4900 RU surface, the curves for **59** and **60** with a 4900 RU surface.

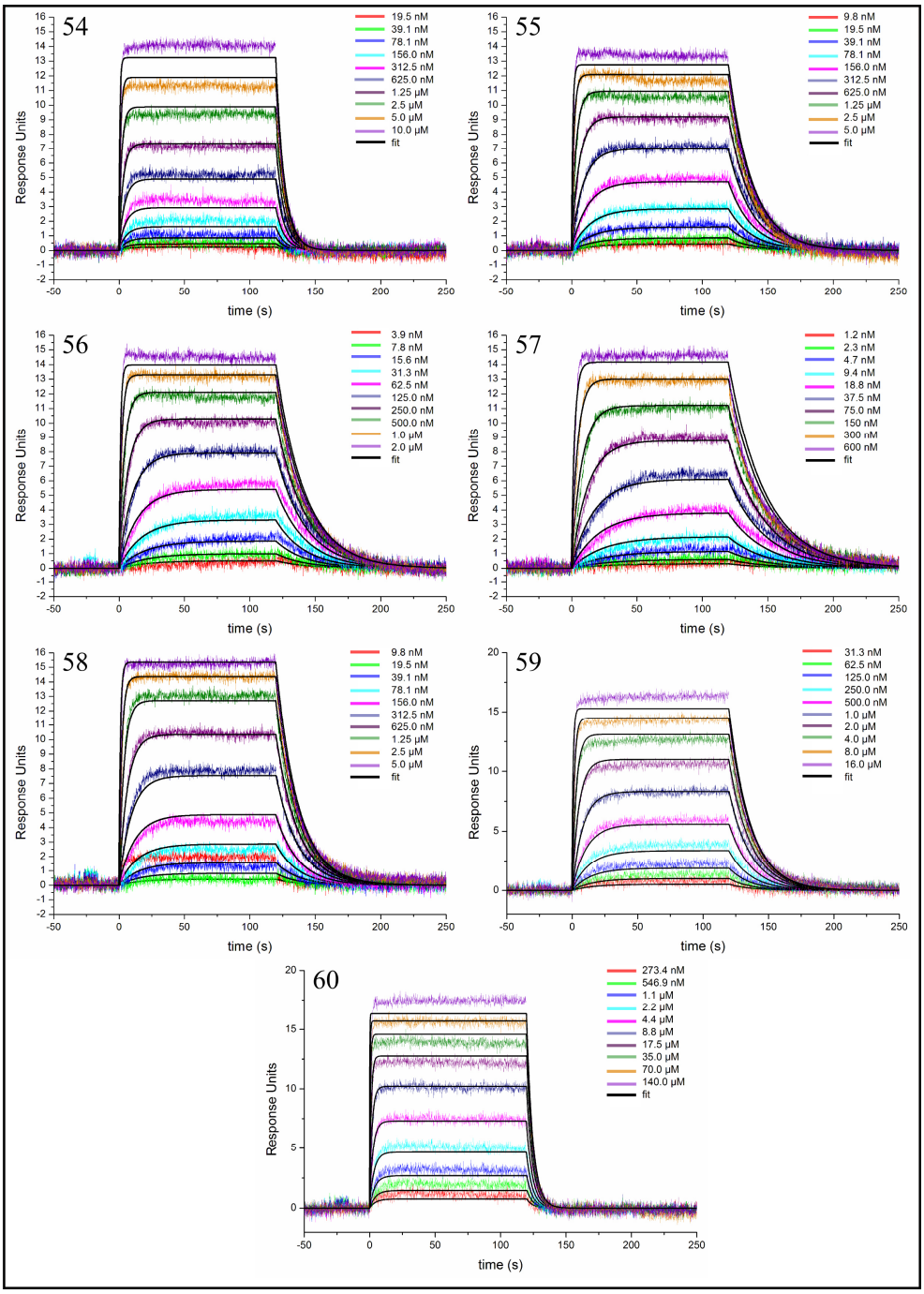



**Figure 10-8 Example sensorgrams for 54-60.** The curves for **54**-**58** were measured with a 4700 RU surface, the curves for **59** and **60** with a 4800 RU surface.

## 10.3 Analytical Data for Ligand 52

### 10.3.1 Ligand 52:

Mp: 232°C .$^{1}$H-NMR ($d_6$-DMSO, 30 °C, 500 MHz): $\delta_H$(ppm) =4.56 (d, $^{3}J$ = 6.3 Hz,2H, H10),4.91 (s, 2H, H7),6.92 (dt, $^{3}J_{H,F}$, $^{3}J_{H,H}$= 8.6 Hz, $^{4}J$ = 2.6 Hz, 1H, H2),7.07(dd, $^{3}J_{H,F}$= 11.0 Hz, $^{4}J$ = 2.6 Hz, 1H, H6),7.44 (pst, $^{3}J$ = 7.6 Hz, 1H, H13),7.58 (d, $^{3}J$ = 8.0 Hz, 1H, H12),7.81 (d, $^{3}J$ = 7.7 Hz, 1H, H14),7.91 (s, 1H, H16),7.93 (dd, $^{3}J$ = 8.7 Hz, $^{4}J_{H,F}$= 7.3 Hz, 1H, H3),9.12 (t, $^{3}J$ = 6.2 Hz, 1H, N*H*),13.10 (bs, 1H, COO*H*).$^{13}$C-NMR ($d_6$-DMSO, 30 °C, 100 MHz): $\delta_C$(ppm) =42.5 (C10),65.9 (7),101.7 (d, $^{2}J_{C,F}$= 26.0 Hz, C6),108.2 (d, $^{2}J_{C,F}$= 22.1 Hz, C2),119.0 (d, $^{4}J_{C,F}$= 2.9 Hz, C4), 127.7, 128.1, 128.5 (C13, C14, C16),130.8 (C15),131.8 (C12),133.0 (d, $^{3}J_{C,F}$= 10.6 Hz, C3), 140.0 (C11),157.2 (d, $^{3}J_{C,F}$= 11.6 Hz, C5),163.7 (C9),164.3 (d, $^{1}J_{C,F}$= 248.5 Hz, C1),167.3 (C17),170.0 (C8).MS (ESI-):*m/z* (%) = 172 (100, [*M*-2H]$^{2-}$), 346 (27, [*M*-H]$^{-}$).HRMS(ESI+) calcd for $C_{17}H_{15}FNO_6$[*M*+H]$^{+}$348.088341; found:348.088200. Anal. calcd for: $C_{17}H_{14}FNO_6$:C, 58.79; H, 4.06; N, 4.03; found:C,58.79; H, 4.23; N, 4.29.

## 10.4 Analytical Data for Ligands 54-60

### 10.4.1 Ligands 54-58:

**Benzenesulfonamide** (**54**) lit.[214-216]:mp: 151-153 °C(lit.[216] 155-158 °C);$^1$H NMR (500 MHz, DMSO-$d_6$) δ 7.85-7.82 (m, 2H), 7.61-7.54 (m, 3H), 7.32 (bs, 2H); MS (ES+) *m/z* 175 $[M + NH_4]^+$; analysis (calcd., found for $C_6H_7NO_2S$): C (45.85, 45.30), H (4.49, 4.43), N (8.91, 9.02).

**4-Methylbenzene-1-sulfonamide** (**55**) lit.[214-216]: mp: 137-139 °C(lit.[216]136 °C); $^1$H NMR (400 MHz, DMSO-$d_6$) δ 7.71 (dm, 2H, *J* = 8.2 Hz), 7.36 (d, 2H, *J* = 8.5 Hz), 7.25 (bs, 2H), 2.37 (s, 3H); MS (ES+) *m/z* 189 $[M + NH_4]^+$; analysis (calcd., found for $C_7H_9NO_2S$): C (49.11, 48.97), H, (5.30, 5.27), N, (8.18, 8.13).

**4-Ethylbenzene-1-sulfonamide** (**56**) lit.[217]:mp: 110-111 °C(lit.[217] 109-111 °C); $^1$H NMR (500 MHz, DMSO-$d_6$) δ 7.73 (dm, 2H, *J* = 8.3 Hz), 7.39 (dm, 2H, *J* = 8.3 Hz), 7.23 (bs, 2H), 2.67 (q, 2H, *J* = 7.7 Hz ); 1.19 (t, 3H, *J* = 7.7 Hz) MS (ES+) *m/z* 203 $[M + NH_4]^+$; analysis (calcd., found for $C_8H_{11}NO_2S$): C (51.87, 51.86), H (5.99, 5.94), N (7.56, 7.72).

**4-Propylbenzene-1-sulfonamide** (**57**) lit.[218]:mp: 101-105 °C(lit.[218] 98-101 °C) $^1$H NMR (400 MHz, DMSO-$d_6$) δ 7.73 (dm, 2H, *J* = 8.2 Hz), 7.37 (dm, 2H, *J* = 8.0 Hz), 7.25 (bs, 2H), 2.62 (t, 2H, *J* = 7.6 Hz ); 1.61 (sext, 2H, *J* = 7.6 Hz), 0.89 (t, 2H, *J* = 7.3 Hz) MS (ES+) *m/z* 217 (100, $[M + NH_4]^+$); analysis (calcd., found for $C_9H_{13}NO_2S$): C (54.25, 53.81), H (6.58, 6.55), N (7.03, 7.21).

**4-(2-Hydroxyethyl)-benzene-1-sulfonamide (58)** lit.[219]:The commercial compound was chromatographed twice on a prepacked silica gel column utilizing chloroform:ethyl acetate, 10:1 (v/v) as eluent. mp: 85-88 °C(lit.[219]54-55 °C)**;**TLC ($CHCl_3$:EtOAc, 10:1 v/v): $R_f$ = 0.13; $^1$H NMR (500 MHz, DMSO-$d_6$) δ 7.72 (dm, 2H, *J* = 8.3 Hz), 7.40 (dm, 2H, *J* = 8.3 Hz), 7.24 (bs, 2H), 4.68 (bt, 1H) 3.63 (m, 2H) 2.79 (t, 2H, *J* = 6.7 Hz); MS (ES+) *m/z* 219 $[M + NH_4]^+$; analysis (calcd., found for $C_9H_{13}NO_2S$): C (47.75, 47.69), H, (5.51,5.67), N, (6.96, 6.88).

### 10.4.2 Ligands 59 and 60:

**4-Carbonyloxybenzene-1-sulfonamide (4-sulfamoylbenzoic acid)** (**59**) lit.[220, 221]: mp: 276-277 °C (lit.[222]282-283 °C); $^1$H NMR (500 MHz, DMSO-$d_6$) δ 13.33 (bs, 1H), 8.11 (dm, 2H, $J$ = 8.6 Hz), 7.94 (dm, 2H, $J$ = 8.6 Hz), 7.51 (bs, 2H); MS (ES-) $m/z$ 200 (100, $[M - H]^-$); analysis (calcd., found for $C_7H_7NO_4S$): C (41.79, 42.14), H (3.51, 3.66), N (6.96, 7.13).

**4-Aminobenzene-1-sulfonamide (sulfanilamide) (60)** lit.[221]: mp: 165-167°C(lit.[222] 163-165 °C); $^1$H NMR (500 MHz, DMSO-$d_6$) δ 7.45 (dm, 2H, $J$ = 8.9 Hz), 6.58 (dm, 2H, $J$ = 8.9 Hz), 5.76 (bs, 2H),MS (ES+) $m/z$ 190 (100, $[M + NH_4]^+$); analysis (calcd., found for $C_6H_8N_2O_2S$): C (41.85, 41.86), H (4.68, 4.69), N (16.26, 16.29).

## 10.5 Calculation of the Purity of Ligand 58 According to qNMR Measurements

**P** = purity of sample (%)
$\mathbf{I_{analyte}}$ = integral of analyte signal
$\mathbf{N_{CRM}}$ = number of CRM protons
$\mathbf{M_{analyte}}$ = molecular mass of analyte
$\mathbf{m_{CRM}}$ = mass of CRM
$\mathbf{I_{CRM}}$ = integral of CRM signal
$\mathbf{N_{analyte}}$ = number of analyte protons
$\mathbf{M_{CRM}}$ = molecular mass of CRM
$\mathbf{m_{sample}}$ = mass of sample/analyte
$\mathbf{P_{CRM}}$ = purity of CRM (%)

qNMR data of commercial **58**:

| | |
|---|---|
| ($\mathbf{I_{analyte}}$) of **58**: | 1.61 |
| ($\mathbf{N_{CRM}}$) of CRM | 2 |
| ($\mathbf{M_{analyte}}$) of **58**: | 201.24 g/mol |
| ($\mathbf{m_{CRM}}$) of CRM: | 4.22 mg |
| ($\mathbf{I_{CRM}}$) of CRM. | 2.00 |
| ($\mathbf{N_{analyte}}$) of **58**: | 2 |
| ($\mathbf{M_{CRM}}$) of CRM: | 116.07 g/mol |
| ($\mathbf{m_{sample}}$) of **58**: | 6.93 mg |
| ($\mathbf{P_{CRM}}$) of CRM: | 100% |

purity (**P**) of commercial **58** in %:

$$\mathbf{P = I_{analyte} * N_{CRM} * M_{analyte} * m_{CRM} * p_{CRM} / I_{CRM} * N_{analyte} * M_{CRM} * m_{sample}}$$

**P** = 1.61* 2 * 201.24 g/mol * 4.22 mg * 100% **/** 2.00 * 2 * 116.07 g/mol * 6.93 mg

= **84.99%**

qNMR data of chromatographed **58** (1. purification step):

($\mathbf{I_{analyte}}$) of **58**: 1.84
($\mathbf{N_{CRM}}$) of CRM 2
($\mathbf{M_{analyte}}$) of **58**: 201.24 g/mol
($\mathbf{m_{CRM}}$) of CRM: 4.16 mg
($\mathbf{I_{CRM}}$) of CRM: 2.00
($\mathbf{N_{analyte}}$) of **58**: 2
($\mathbf{M_{CRM}}$) CRM: 116.07 g/mol
($\mathbf{m_{sample}}$) of **58**: 7.04 mg
($\mathbf{P_{CRM}}$) of CRM: 100%

purity (**P**) of chromatographed **58** in % (1. purification step):

**P** = 1.84 * 2 * 201.24 g/mol * 4.16 mg * 100% **/** 2.00 * 2 * 116.07 g/mol * 7.04 mg

= **94.25%**

qNMR data of chromatographed **58** (2. purification step):

($\mathbf{I_{analyte}}$) of **58**: 2.07
($\mathbf{N_{CRM}}$) of CRM 2
($\mathbf{M_{analyte}}$) of **58**: 201.24 g/mol
($\mathbf{m_{CRM}}$) of CRM: 4.07 mg
($\mathbf{I_{CRM}}$) of CRM. 2.00
($\mathbf{N_{analyte}}$) **58** 2
($\mathbf{M_{CRM}}$) of CRM: 116.07 g/mol
($\mathbf{m_{sample}}$) of **58**: 7.48 mg
($\mathbf{P_{CRM}}$) of CRM: 100%

purity (**P**) of chromatographed **58** in % (2. purification step):

**P** = 2.07 * 2 * 201.24 g/mol * 4.07 mg * 100% **/** 2.00 * 2 * 116.07 g/mol * 7.48 mg

= **97.64%**

## 10.6 Crystallographic Table of ALR2 Complex Crystal Structures

**Table 10-3 Crystallographic table of ALR2 complex crystal structures.**

| | ALR2-**13** | ALR2-**28**-$P2_1$ | ALR2-**30** | ALR2-**31** | ALR2-**37** |
|---|---|---|---|---|---|
| **PDB entry** | **4QX4** | **4YU1** | **4PR4** | **4PRR** | **4NKC** |
| BESSY beamline | 14.2 | 14.1 | 14.2 | 14.2 | 14.2 |
| **Data collection and processing** | | | | | |
| Wavelength [Å] | 0.91841 | 0.91841 | 0.91841 | 0.91841 | 0.91841 |
| Space group | $P2_1$ | $P2_1$ | $P2_1$ | $P2_1$ | $P2_1$ |
| **Unit cell parameters** | | | | | |
| a,b,c [Å] | 49.4, 66.9, 47.4 | 49.5, 66.4, 47.4 | 49.4, 66.5, 47.3 | 49.3, 66.9, 47.2 | 49.2, 66.4, 47.2 |
| β [°] | 91.9 | 92.3 | 92.3 | 92.5 | 92.5 |
| Matthews coefficient [$Å^3$/Da]/ Solvent content [%] | 2.3/ 45.6 | 2.2/ 42.9 | 2.2/ 42.8 | 2.2/ 42.8 | 2.1/ 42.3 |
| **Diffraction data** [A] | | | | | |
| Resolution range [Å] | 17.84-1.26 (1.28-1.26) | 47.39-1.02 (1.08-1.02) | 49.33-1.06 (1.08-1.06) | 18.78-1.01 (1.03-1.01) | 18.46-1.12 (1.14-1.12) |
| Unique reflections | 79172(3759) | 150587(23315) | 129820(6126) | 148147(7006) | 116660(5682) |
| R(I)sym [%] | 5.1 (30.0) | 4.9 (46.3) | 4.1 (42.7) | 5.6 (42.1) | 5.2 (40.4) |
| Completeness [%] | 95.0 (90.4) | 96.7 (92.9) | 93.7 (88.5) | 92.6 (87.2) | 99.2 (99.2) |
| Redundancy | 2.7 (2.7) | 3.4 (3.2) | 2.6 (2.5) | 2.1 (1.8) | 2.9 (2.7) |
| I/σ(I) | 21.6 (3.9) | 14.0 (2.5) | 24.5 (2.5) | 16.9 (2.0) | 18.9 (2.4) |
| **Refinement** | | | | | |
| Resolution range [Å] | 17.84-1.26 | 47.39-1.02 | 49.33-1.06 | 18.78-1.01 | 18.46-1.12 |
| Reflections used in refinement (work/free) | 79143 (75204/3939) | 150569 (143039/7530) | 129790 (123246/6544) | 148117 (140709/7408) | 116632 (110758/5874) |
| Final R values for all reflections (work/free) [%] | 13.3/15.3 | 13.3/15.0 | 13.7/15.3 | 14.4/15.5 | 12.2/14.2 |
| Protein residues | 316 | 316 | 315 | 315 | 314 |
| Atoms inhibitor ($1^{st}/2^{nd}$), cofactor | 18/18/48 | 19/-/48 | 20/-/48 | 19/-/48 | 20/-/48 |
| Water molecules | 334 | 413 | 359 | 334 | 360 |
| **RMSDs from ideality** [B] | | | | | |
| Bond lengths[Å]/ bond angles [°] | 0.006/1.2 | 0.008/1.4 | 0.006/1.3 | 0.006/1.3 | 0.007/1.4 |
| **Ramachandran plot** [C] | | | | | |
| Residues in most favored regions [%] | 91.0 | 89.9 | 91.7 | 91.7 | 91.7 |
| Residues in additional allowed regions [%] | 9.0 | 10.1 | 8.3 | 8.3 | 8.3 |
| Residues in generously allowed regions [%] | 0 | 0 | 0 | 0 | 0 |
| **Mean B-factor [$Å^2$]** [D] | | | | | |
| Protein | 11.2 | 9.8 | 10.6 | 10.4 | 12.3 |
| Inhibitor (1st/2nd), Cofactor | 9.3/18.0/6.9 | 8.6/-/6.2 | 12.2/-/6.4 | 10.1/-/6.7 | 13.3/-/6.9 |
| Water molecules | 24.2 | 23.2 | 24.1 | 23.3 | 25.6 |

**Continuation of Table 10-3.**

| | ALR2-**43** | ALR2-**47** | ALR2-**48** | ALR2-**49** | ALR2-**50** |
|---|---|---|---|---|---|
| **PDB entry** | **4PRT** | **4PUU** | **4PUW** | **4Q7B** | **4QR6** |
| BESSY beamline | 14.1 | 14.2 | 14.2 | 14.2 | 14.2 |
| **Data collection and processing** | | | | | |
| Wavelength [Å] | 0.91841 | 0.91841 | 0.91841 | 0.91841 | 0.91841 |
| Space group | $P2_1$ | $P2_1$ | $P2_1$ | $P2_1$ | $P2_1$ |
| **Unit cell parameters** | | | | | |
| a,b,c [Å] | 49.3, 67.0, 47.3 | 49.3, 66.4, 47.2 | 49.4, 66.7, 47.4 | 49.3, 66.7, 47.3 | 49.5, 66.7, 47.4 |
| β [°] | 92.5 | 91.8 | 92.0 | 92.1 | 92.6 |
| Matthews coefficient [$Å^3$/Da]/ Solvent content [%] | 2.2/ 43.1 | 2.1/ 42.4 | 2.2/ 43.1 | 2.2/ 42.8 | 2.2/ 43.1 |
| **Diffraction data** [A] | | | | | |
| Resolution range [Å] | 38.6-0.96 (1.02-0.96) | 17.32-1.14 (1.16-1.14) | 49.38-1.12 (1.14-1.12) | 18.73-1.19 (1.21-1.19) | 49.47-1.05 (1.07-1.05) |
| Unique reflections | 185561(29443) | 109695(5447) | 117148(5781) | 97712(4840) | 142734(7043) |
| R(I)sym [%] | 3.4 (40.9) | 5.5 (36.9) | 5.9 (42.6) | 5.5 (44.6) | 4.3 (46.6) |
| Completeness [%] | 99.3 (97.6) | 99.1 (99.0) | 99.7 (100.0) | 99.8 (98.9) | 100.0 (100.0) |
| Redundancy | 3.5 (3.1) | 2.9 (2.6) | 2.5 (2.3) | 3.2 (3.0) | 3.9 (3.6) |
| I/σ(I) | 19.4 (2.8) | 19.1 (2.7) | 16.0 (2.4) | 20.0 (2.5) | 30.1 (2.8) |
| **Refinement** | | | | | |
| Resolution range [Å] | 38.6-0.96 | 17.32-1.14 | 49.38-1.12 | 18.73-1.19 | 49.47-1.05 |
| Reflections used in refinement (work/free) | 185561 (176283/9278) | 109668 (104171/5497) | 117119 (111252/5867) | 97684 (92805/4879) | 142704 (135543/7161) |
| Final R values for all reflections (work/free) [%] | 12.6/13.7 | 13.7/15.6 | 13.7/15.7 | 13.8/16.0 | 14.1/15.7 |
| Protein residues | 316 | 316 | 316 | 315 | 316 |
| Atoms Inhibitor | 24 | 15 | 18 | 22 | 25 |
| Water molecules | 412 | 380 | 402 | 370 | 355 |
| **RMSDs from ideality** [B] | | | | | |
| Bond lengths[Å]/ bond angles [°] | 0.006/1.3 | 0.005/1.2 | 0.006/1.3 | 0.006/1.2 | 0.005/1.3 |
| **Ramachandran plot** [C] | | | | | |
| Residues in most favored regions [%] | 91.7 | 91.4 | 91.0 | 91.3 | 91.7 |
| Residues in additional allowed regions [%] | 8.3 | 8.6 | 9.0 | 8.7 | 8.3 |
| Residues in generously allowed regions [%] | 0 | 0 | 0 | 0 | 0 |
| **Mean B-factor [$Å^2$]** [D] | | | | | |
| Protein | 9.3 | 11.0 | 9.8 | 11.1 | 10.7 |
| Inhibitor , Cofactor | 7.0 (-), 5.9 | 10.7 (1.0), 6.1 | 9.4 (1.0), 6.2 | 11.3 (1.0), 6.4 | 8.5 (0.8), 6.4 |
| Ala 299- Tyr309 | - | 21.4 (1.9) | 12.7 (1.3) | 22.5 (2.0) | 14.5 (1.4) |
| Water molecules | 24.3 | 24.2 | 22.3 | 24.9 | 22.3 |

**Continuation of Table 10-3.**

| | ALR2-**51** | ALR2-**52** | ALR2-**53** |
|---|---|---|---|
| **PDB entry** | **4QBX** | **4YS1** | **4RPQ** |
| BESSY beamline | 14.2 | 14.2 | 14.2 |
| **Data collection and processing** | | | |
| Wavelength [Å] | 0.91841 | 0.91841 | 0.91841 |
| Space group | $P2_1$ | $P2_1$ | $P2_1$ |
| **Unit cell parameters** | | | |
| a,b,c [Å] | 49.4, 66.9, 47.4 | 49.3, 66.6, 47.3 | 49.6, 66.8, 47.2 |
| β [°] | 92.3 | 92.1 | 93.1 |
| Matthews coefficient [$Å^3$/Da]/ Solvent content [%] | 2.2/ 43.2 | 2.2/ 42.7 | 2.2/ 43.1 |
| **Diffraction data**[A] | | | |
| Resolution range [Å] | 17.43-0.98 (1.00-0.98) | 49.28-1.07 (1.13-1.07) | 35.08-1.20 (1.22-1.20) |
| Unique reflections | 175882(8807) | 131957(20414) | 92652(4516) |
| R(I)sym [%] | 6.0 (41.4) | 6.5 (46.6) | 3.9 (41.8) |
| Completeness [%] | 100.0 (99.9) | 97.6 (93.6) | 96.8 (94.2) |
| Redundancy | 3.7 (3.5) | 3.8 (3.6) | 2.4 (2.3) |
| I/σ(I) | 22.1 (2.6) | 14.0 (2.7) | 24.1 (2.5) |
| **Refinement** | | | |
| Resolution range [Å] | 17.43-0.98 | 49.28-1.07 | 35.08-1.20 |
| Reflections used in refinement (work/free) | 175851 (167049/8802) | 131957 (125359/6598) | 92626 (87981/4645) |
| Final R values for all reflections (work/free) [%] | 13.5/15.1 | 13.2/14.8 | 14.2/16.5 |
| Protein residues | 316 | 316 | 314 |
| Atoms Inhibitor | 25 | 27 | 30 |
| Water molecules | 396 | 411 | 293 |
| **RMSDs from ideality** [B] | | | |
| Bond lengths[Å]/ bond angles [°] | 0.006/1.3 | 0.005/1.3 | 0.005/1.2 |
| **Ramachandran plot** [C] | | | |
| Residues in most favored regions [%] | 91.4 | 90.6 | 92.0 |
| Residues in additional allowed regions [%] | 8.6 | 9.4 | 8.0 |
| Residues in generously allowed regions [%] | 0 | 0 | 0 |
| **Mean B-factor [$Å^2$]** [D] | | | |
| Protein | 8.8 | 9.0 | 12.2 |
| Inhibitor , Cofactor | 10.5 (1.2) | 15.3 (1.7) | 15.0 (1.2) |
| Ala 299- Tyr309 | 7.3 (0.8), 5.1 | 8.8 (1.0), 5.2 | 10.7 (0.9), 7.9 |
| Water molecules | 22.5 | 23.1 | 24.3 |

## 10.7 Crystallographic Table of hCAII Complex Crystal Structures

**Table 10-4 Crystallographic table of hCAII complex crystal structures**

| | hCAII-**54** | hCAII-**55** | hCAII-**56** | hCAII-**57** | hCAII-**58** |
|---|---|---|---|---|---|
| **PDB entry** | **4YX4** | **4YXI** | **4YXO** | **4YXU** | **4YYT** |
| BESSY beamline | 14.1 | 14.1 | 14.2 | 14.1 | 14.1 |
| **Data collection and processing** | | | | | |
| Wavelength [Å] | 0.91841 | 0.91841 | 0.91841 | 0.91841 | 0.91841 |
| Space group | $P2_1$ | $P2_1$ | $P2_1$ | $P2_1$ | $P2_1$ |
| **Unit cell parameters** | | | | | |
| a,b,c [Å] | 42.4, 41.5, 72.3 | 42.5, 41.5, 72.5 | 42.4, 41.6, 72.3 | 42.3, 41.7, 72.2 | 42.5, 41.7, 72.3 |
| β [°] | 104.4 | 104.5 | 104.3 | 104.3 | 104.3 |
| Matthews coefficient [$Å^3$/Da]/ Solvent content [%] | 2.1/ 40.6 | 2.1/ 40.9 | 2.1/ 40.8 | 2.1/ 40.8 | 2.1/ 41.0 |
| **Diffraction data** [A] | | | | | |
| Resolution range [Å] | 35.70-1.01 (1.07-1.01) | 41.11-0.96 (1.02-0.96) | 35.05-1.06 (1.12-1.06) | 41.44-1.08 (1.15-1.08) | 41.14-1.07 (1.13-1.07) |
| Unique reflections | 126327(19756) | 148298(23592) | 108590(17145) | 103099(16077) | 107591(17203) |
| R(I)sym [%] | 4.9 (41.9) | 4.4 (38.4) | 7.7 (49.0) | 5.5 (38.6) | 5.8 (49.1) |
| Completeness [%] | 98.7 (96.0) | 99.2 (98.1) | 98.0 (96.3) | 98.0 (95.0) | 99.3 (98.9) |
| Redundancy | 3.2 (2.8) | 3.5 (3.2) | 3.6 (3.6) | 3.3 (3.2) | 3.2 (3.0) |
| I/σ(I) | 11.0 (2.0) | 13.3 (2.4) | 9.1 (2.1) | 12.1 (2.6) | 11.6 (2.1) |
| **Refinement** | | | | | |
| Resolution range [Å] | 35.70-1.01 | 41.11-0.96 | 35.05-1.06 | 41.44-1.08 | 41.44-1.08 |
| Reflections used in refinement (work/free) | 126327 (120010/6317) | 148294 (140879/7415) | 108590 (103160/5430) | 103099 ( 97944/5155) | 107591 (102211/5380) |
| Final R values for all reflections (work/free) [%] | 15.1/17.0 | 14.0/15.3 | 14.9/16.4 | 14.4/16.3 | 14.4/16.6 |
| Protein residues | 257 | 257 | 257 | 258 | 257 |
| Atoms Inhibitor | 10/-/-/- | 11/11/-/- | 12/12/-/- | 13/13/-/- | 13/13/13/11 |
| Water molecules | 249 | 258 | 235 | 244 | 245 |
| **RMSDs from ideality** [B] | | | | | |
| Bond lengths[Å]/ bond angles [°] | 0.005/1.1 | 0.005/1.1 | 0.005/1.1 | 0.005/1.1 | 0.005/1.1 |
| **Ramachandran plot** [C] | | | | | |
| Residues in most favored regions [%] | 88.4 | 88.9 | 89.4 | 88.9 | 88.9 |
| Residues in additional allowed regions [%] | 11.1 | 10.6 | 10.2 | 10.6 | 10.6 |
| Residues in generously allowed regions [%] | 0.5 | 0.5 | 0.5 | 0.5 | 0.5 |
| **Mean B-factor [$Å^2$]** [D] | | | | | |
| Protein | 12.6 | 10.7 | 12.5 | 10.4 | 10.4 |
| Inhibitor ($1^{st}$, $2^{nd}$, $3^{rd}$, $4^{th}$) | 10.7/-/-/- | 8.4/16.1/-/- | 11.1/19.1/-/- | 9.5/14.6/-/- | 8.3/13.6/19.0/17.4 |
| Water molecules | 27.0 | 25.0 | 26.6 | 24.0 | 24.5 |

**Table 10-3 and Table 10-4:**
[A]Numbers in brackets represent highest resolution shells.
[B]RMSD were calculated with MolProbity[207] as implemented in the Phenix suite.[141]
[C]PROCHECK[213] was used to generate the Ramachandran plots.
[D]MOLEMAN[223] was used to calculate the mean *B*-factors. Numbers in brackets represent the ratios of the mean *B*-factor of the residues 299-309 or the ligand, respectively, compared to the mean *B*-factor of all protein residues.

## 10.8 Sequencing Results of ALR2 and hCAII

```
CLUSTAL format alignment by MAFFT (v7.245)

ALR2            FCLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASRILLNNGAKMPILGLGTWKSPPGQVTE
UNIPROT_P15121  ------------------------------ASRLLLNNGAKMPILGLGTWKSPPGQVTE
                                              ***:*************************

ALR2            AVKVAIDVGYRHIDCAHVYQNENEVGVAIQEKLREQVVKREELFIVSKLWCTYHEKGLVK
UNIPROT_P15121  AVKVAIDVGYRHIDCAHVYQNENEVGVAIQEKLREQVVKREELFIVSKLWCTYHEKGLVK
                ************************************************************

ALR2            GACQKTLSDLKLDYLDLYLIHWPTGFKPGKEFFPLDESGNVVPSDTNILDTWAAMEELVD
UNIPROT_P15121  GACQKTLSDLKLDYLDLYLIHWPTGFKPGKEFFPLDESGNVVPSDTNILDTWAAMEELVD
                ************************************************************

ALR2            EGLVKAIGISNFNHLQVEMILNKPGLKYKPAVNQIECHPYLTQEKLIQYCQSKGIVVTAY
UNIPROT_P15121  EGLVKAIGISNFNHLQVEMILNKPGLKYKPAVNQIECHPYLTQEKLIQYCQSKGIVVTAY
                ************************************************************

ALR2            SPLGSPDRPWAKPEDPSLLEDPRIKAIAAKHNKTTAQVLIRFPMQRNLVVIPKSVTPERI
UNIPROT_P15121  SPLGSPDRPWAKPEDPSLLEDPRIKAIAAKHNKTTAQVLIRFPMQRNLVVIPKSVTPERI
                ************************************************************

ALR2            AENFKVFDFELSSQDMTTLLSYNRNWRVCALLSCTSHKDYPFHEEF
UNIPROT_P15121  AENFKVFDFELSSQDMTTLLSYNRNWRVCALLSCTSHKDYPFHEEF
                *********************************************
```

**Figure 10-9 Sequencing result of ALR2 and comparison with the UniProt entry P15121.** The sequence conflict in position 4 (L→I) is already documented in the UniProt entry.

```
CLUSTAL format alignment by MAFFT (v7.245)

hCAII           MSHHWGYGKHNGPEHWHKDFPIAKGERQSPVDIDTHTAKYDPSLKPLSVSYDQATSLRIL
UNIPROT_P00918  -SHHWGYGKHNGPEHWHKDFPIAKGERQSPVDIDTHTAKYDPSLKPLSVSYDQATSLRIL
                 ************************************************************

hCAII           NNGHAFNVEFDDSQDKAVLKGGPLDGTYRLIQFHFHWGSLDGQGSEHTVDKKKYAAELHL
UNIPROT_P00918  NNGHAFNVEFDDSQDKAVLKGGPLDGTYRLIQFHFHWGSLDGQGSEHTVDKKKYAAELHL
                ************************************************************

hCAII           VHWNTKYGDFGKAVQQPDGLAVLGIFLKVGSAKPGLQKVVDVLDSIKTKGKSADFTNFDP
UNIPROT_P00918  VHWNTKYGDFGKAVQQPDGLAVLGIFLKVGSAKPGLQKVVDVLDSIKTKGKSADFTNFDP
                ************************************************************

hCAII           RGLLPESLDYWTYPGSLTTPPLLECVTWIVLKEPISVSSEQVLKFRKLNFNGEGEPEELM
UNIPROT_P00918  RGLLPESLDYWTYPGSLTTPPLLECVTWIVLKEPISVSSEQVLKFRKLNFNGEGEPEELM
                ************************************************************

hCAII           VDNWRPAQPLKNRQIKASFK
UNIPROT_P00918  VDNWRPAQPLKNRQIKASFK
                ********************
```

**Figure 10-10 Sequencing result of hCAII and comparison with the UniProt entry P00918.**

# Danksagung

Ich danke ganz besonders *Herrn Prof. Dr. Gerhard Klebe* für die sehr gute Betreuung und das interessante Promotionsthema. Aus unseren gemeinsamen Diskussionen gingen viele Anregungen hervor, die meinen Projekten oft den entscheidenden Impuls gaben. Ich möchte mich dafür bedanken, dass ich immer die Freiheit hatte, eigene Ideen zu verfolgen. Die Möglichkeit, an verschiedenen Konferenzen teilzunehmen, hat sich ebenso als sehr vorteilhaft erwiesen wie das Eingehen von mehreren Kooperationen zu deutschen und internationalen Arbeitsgruppen. Nur so war es möglich, die hier beschriebenen Untersuchungen durchzuführen.

*Herrn Prof. Dr. Andreas Heine* danke ich für die Ausbildung in der Proteinkristallographie und für unzählige Stunden, in denen wir kristallographische Probleme diskutierten. Mit seiner großen Erfahrung konnte er mir immer weiterhelfen, wenn einmal wieder unvorhergesehene Probleme auftraten.

Ich möchte mich bei *Prof. Dr. Martin Schlitzer* und seiner Arbeitsgruppe, besonders *Frau Dr. Regina Ortmann*, für die Synthese einer der Aldose Reduktase-Inhibitor-Serien bedanken. Auch danke ich *Herrn Patrick Mäder* für seine Unterstützung.

*Frau Prof. Dr. Wibke Diederich* und ihrer Arbeitsgruppe, besonders *Dr. Frithjof Scheer* und *Dr. Philipp Toth*, danke ich für die Bereitstellung der IDD-Liganden-Serie. Besonders hervorheben möchte ich die schnelle Durchführung der Synthesen, die es uns erlaubte nach jeder Synthese-Runde die richtigen Liganden zu identifizieren, um die Spezifitätstasche der ALR2 näher zu untersuchen.

Ich danke *Frau Alexandra Cousido-Siah, Herrn Dr. Francesc Xavier Ruiz* und *Herrn Prof. Dr. Alberto Podjarny* (CNRS, Illkirch, France) für die gute Zusammenarbeit und für den wissenschaftlichen Austausch.

Ich möchte mich bei *Dr. Roberto Gaspari* und *Prof. Dr. Andrea Cavalli* (Italian Institute of Technology, Genova, Italy) für die Kooperation auf dem Carboanhydrase-Projekt bedanken. Außerdem danke ich *Dr. Daniel Schwarz* und *Dr. Jörg Bomke* (Merck Serono Research,

Darmstadt) für die großartige Unterstützung und für die Möglichkeit, die SPR-Messungen in ihrem Labor durchführen zu können. *Hans-Dieter Gerber* danke ich für seine Mithilfe, besonders für die Aufreinigung eines der Inhibitoren, ohne die dieses Projekt nicht möglich geworden wäre.

*Herrn Dr. Milan Stefek* (Bratislava) danke ich für die erfolgreiche Kooperation auf dem Gebiet der ALR2-Inhibitoren.

Meinen Vorgängern auf dem hCAII- und ALR2-Projekten möchte ich für die überlassenen Protokolle danken. Besonders möchte ich *Herrn Dr. Johannes Schulze Wischeler* für seine sehr gute Einweisung in das hCAII-Projekt und *Herrn Dr. Holger Steuber* und *Frau Dr. Cornelia Koch* für die guten Tipps bezüglich des ALR2-Projektes danken. *Herrn Dr. Christian Hasewinkel* und *Stefanie Dörr* danke ich für viele gute Hinweise im S1-Labor.

*Herr Dr. Tobias Craan* und *Herr Dr. Florian Immekus* haben mich gerade zu Beginn meiner Arbeit sehr unterstützt. Ich danke ihnen für die tolle Einweisung in die Wissenschaft- und in alles, was sonst noch wichtig ist in der AG Klebe.

Ich danke *Frau Lydia Hartleben* für ihre Unterstützung in allen organisatorischen Dingen.

*Herrn Christian Sohn* danke ich für seine Hilfe bei dem Betrieb des Röntgengerätes und dafür, dass er bei allen technischen Problemen immer weiß, was zu tun ist.

Ich danke *Michael, Felix T., Phong, Timo, Denis, Thomas, Felix G., Gerd* und *Sven* für die Administration des Computernetzwerkes.

Ich danke meinen Vertiefungsstudenten *Chelsey Ann Collins*, *Philipp Seiler, Lydia Gebauer* und *Hauke Löcken* für ihre wertvolle Mitarbeit.

Ich danke ganz besonders *Chelsey* für all ihre Unterstützung, für ihre Gesellschaft an unzähligen Abenden an der Äkta und für das Korrekturlesen dieser Arbeit.

*Florian* und *Phong* möchte ich für die gute Büroatmosphäre danken.

Desweiteren danke ich allen *Mitgliedern der AG Klebe*, besonders danke ich *Frederik, Phong, Martin, Nicole, Stefan, Florian, Michael* und *Kan* für die schöne Zeit.

# Bibliography

(1) Klebe, G. *Wirkstoffdesign- Entwurf und Wirkung von Arzneistoffen* (2009), 2 ed., Spektrum Akademischer Verlag, Heidelberg.

(2) Imming, P., Sinning, C. and Meyer, A. (2006) Drugs, Their Targets and the Nature and Number of Drug Targets, *Nat. Rev. Drug Discovery 5*, 821-834.

(3) Morgan, S., Grootendorst, P., Lexchin, J., Cunningham, C. and Greyson, D. (2011) The Cost of Drug Development: A Systematic Review, *Health Policy 100*, 4-17.

(4) Ladbury, J. E., Klebe, G. and Freire, E. (2010) Adding Calorimetric Data to Decision Making in Lead Discovery: A Hot Tip, *Nat. Rev. Drug Discovery 9*, 23-27.

(5) Wiseman, T., Williston, S., Brandts, J. F. and Lin, L.-N. (1989) Rapid Measurement of Binding Constants and Heats of Binding Using a New Titration Calorimeter, *Anal. Biochem. 179*, 131-137.

(6) Olsson, T. S. G., Williams, M. A., Pitt, W. R. and Ladbury, J. E. (2008) The Thermodynamics of Protein–Ligand Interaction and Solvation: Insights for Ligand Design, *J. Mol. Biol. 384*, 1002-1017.

(7) Klebe, G. (2015) Applying Thermodynamic Profiling in Lead Finding and Optimization, *Nat. Rev. Drug Discovery 14*, 95-110.

(8) Steuber, H., Heine, A. and Klebe, G. (2007) Structural and Thermodynamic Study on Aldose Reductase: Nitro-Substituted Inhibitors with Strong Enthalpic Binding Contribution, *J. Mol. Biol. 368*, 618-638.

(9) www.sprpages.nl. *last access: 14.07.2015*.

(10) Biacore/ GE Healtcare. *Biacore Sensor Surface Handbook BR-1005-71 Edition AB*, 36- 41.

(11) Fägerstam, L. G., Frostell-Karlsson, Å., Karlsson, R., Persson, B. and Rönnberg, I. (1992) Biospecific Interaction Analysis Using Surface Plasmon Resonance Detection Applied to Kinetic, Binding Site and Concentration Analysis, *J. Chromatogr. 597*, 397-410.

(12) Lu, H. and Tonge, P. J. (2010) Drug-Target Residence Time: Critical Information for Lead Optimization, *Curr. Opin. Chem. Biol. 14*, 467-474.

(13) *Definition and Diagnosis of Diabetes Mellitus and Intermediate Hyperglycemia* (2006), World Health Organization, Geneva, Switzerland.

(14) Maggiore, W. A. (2013) Highs & Lows: Recognizing & Treating Hypoglycemia, Hyperglycemia & Other Diabetes-Related Health Problems, *JEMS 38*, 44-47.

(15) American Diabetes Association. (2015) Classification and Diagnosis of Diabetes, *Diabetes Care 38 (Suppl. 1)*, S8-S16.

(16) American Diabetes Association. (2014) Diagnosis and Classification of Diabetes Mellitus, *Diabetes Care 37 (Suppl 1)*, S81-S90.

(17) Guariguata, L., Whiting, D. R., Hambleton, I., Beagley, J., Linnenkamp, U. and Shaw, J. E. (2014) Global Estimates of Diabetes Prevalence for 2013 and Projections for 2035, *Diabetes Res. Clin. Pract. 103*, 137-149.

(18) American Diabetes Association. (2015) Approaches to Glycemic Treatment, *Diabetes Care 38 Suppl*, S41-S48.

(19) Inzucchi, S. E., Bergenstal, R. M., Buse, J. B., Diamant, M., Ferrannini, E., Nauck, M., Peters, A. L., Tsapas, A., Wender, R. and Matthews, D. R. (2015) Management of Hyperglycemia in Type 2 Diabetes, 2015: A Patient-Centered Approach: Update to a Position Statement of the American Diabetes Association and the European Association for the Study of Diabetes, *Diabetes Care 38*, 140-149.

(20) Brownlee, M. (2001) Biochemistry and Molecular Cell Biology of Diabetic Complications, *Nature 414*, 813-820.

(21) Brownlee, M. (2005) The Pathobiology of Diabetic Complications: A Unifying Mechanism, *Diabetes 54*, 1615-1625.

(22) Pirart, J. (1977) Diabetes Mellitus and Its Degenerative Complications: A Prospective Study of 4,400 Patients Observed between 1947 and 1973 (3rd and Last Part) *Diabete Metab 3*, 245-256.

(23) American Diabetes Association. (2015) Microvascular Complications and Foot Care, *Diabetes Care 38(Suppl. 1)*, S58–S66.

(24) Menz, H. B., Lord, S. R., St George, R. and Fitzpatrick, R. C. (2004) Walking Stability and Sensorimotor Function in Older People with Diabetic Peripheral Neuropathy, *Arch. Phys. Med. Rehabil. 85*, 245-252.

(25) Harris, M. I., Klein, R., Welborn, T. A. and Knuiman, M. W. (1992) Onset of NIDDM Occurs at Least 4-7 Years before Clinical Diagnosis, *Diabetes Care 15*, 815-819.

(26) Spijkerman, A. M., Dekker, J. M., Nijpels, G., Adriaanse, M. C., Kostense, P. J., Ruwaard, D., Stehouwer, C. D., Bouter, L. M. and Heine, R. J. (2003) Microvascular Complications at Time of Diagnosis of Type 2 Diabetes Are Similar among Diabetic Patients Detected by Targeted Screening and Patients Newly Diagnosed in General Practice: The Hoorn Screening Study, *Diabetes Care 26*, 2604-2608.

(27) Kharroubi, A. T. and Darwish, H. M. (2015) Diabetes Mellitus: The Epidemic of the Century, *World J. Diabetes 6*, 850-867.

(28) Gabbay, K. H., Merola, L. O. and Field, R. A. (1966) Sorbitol Pathway: Presence in Nerve and Cord with Substrate Accumulation in Diabetes, *Science 151*, 209-210.

(29) Engerman, R. L., Kern, T. S. and Larson, M. E. (1994) Nerve Conduction and Aldose Reductase Inhibition During 5 Years of Diabetes or Galactosaemia in Dogs, *Diabetologia 37*, 141-144.

(30) Chalk, C., Benstead, T. J. and Moore, F. (2007) Aldose Reductase Inhibitors for the Treatment of Diabetic Polyneuropathy, *Cochrane Database Syst Rev*, CD004572.

(31) Schemmel, K. E., Padiyara, R. S. and D'Souza, J. J. (2010) Aldose Reductase Inhibitors in the Treatment of Diabetic Peripheral Neuropathy: A Review, *J. Diabetes Complications 24*, 354-360.

(32) Yasuda, H., Sonobe, M., Hisanaga, T., Kawabata, T., Maeda, K., Kikkawa, R. and Shigeta, Y. (1992) A Combination of the Aldose Reductase Inhibitor, Statil, and the Prostaglandin E1 Analogue, OP1206.ACD, Completely Improves Sciatic Motor Nerve Conduction Velocity in Streptozocin-Induced Chronically Diabetic Rats, *Metabolism 41*, 778-782.

(33) Wilson, D. K., Bohren, K. M., Gabbay, K. H. and Quiocho, F. A. (1992) An Unlikely Sugar Substrate Site in the 1.65 Å Structure of the Human Aldose Reductase Holoenzyme Implicated in Diabetic Complications, *Science 257*, 81-84.

(34) Biadene, M., Hazemann, I., Cousido, A., Ginell, S., Joachimiak, A., Sheldrick, G. M., Podjarny, A. and Schneider, T. R. (2007) The Atomic Resolution Structure of Human Aldose Reductase Reveals That Rearrangement of a Bound Ligand Allows the Opening of the Safety-Belt Loop, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 63*, 665-672.

(35) Bohren, K. M., Brownlee, J. M., Milne, A. C., Gabbay, K. H. and Harrison, D. H. (2005) The Structure of Apo R268A Human Aldose Reductase: Hinges and Latches That Control the Kinetic Mechanism, *Biochim. Biophys. Acta 1748*, 201-212.

(36) *The Pymol Molecular Graphics System, Version 1.7.4 Schrödinger, Llc.*, New York, NY (USA).

(37) Sotriffer, C. A., Kramer, O. and Klebe, G. (2004) Probing Flexibility and "Induced-Fit" Phenomena in Aldose Reductase by Comparative Crystal Structure Analysis and Molecular Dynamics Simulations, *Proteins: Struct., Funct., Bioinf. 56*, 52-66.

(38) Harrison, D. H., Bohren, K. M., Ringe, D., Petsko, G. A. and Gabbay, K. H. (1994) An Anion Binding Site in Human Aldose Reductase: Mechanistic Implications for the Binding of Citrate, Cacodylate, and Glucose 6-Phosphate, *Biochemistry 33*, 2011-2020.

(39) Steuber, H., Heine, A., Podjarny, A. and Klebe, G. (2008) Merging the Binding Sites of Aldose and Aldehyde Reductase for Detection of Inhibitor Selectivity-Determining Features, *J. Mol. Biol. 379*, 991-1016.

(40) Howard, E. I., Sanishvili, R., Cachau, R. E., Mitschler, A., Chevrier, B., Barth, P., Lamour, V., Van Zandt, M., Sibley, E., Bon, C., Moras, D., Schneider, T. R., Joachimiak, A. and Podjarny, A. (2004) Ultrahigh Resolution Drug Design I: Details of Interactions in Human Aldose Reductase-Inhibitor Complex at 0.66 Å, *Proteins: Struct., Funct., Bioinf. 55*, 792-804.

(41) Koch, C., Heine, A. and Klebe, G. (2011) Tracing the Detail: How Mutations Affect Binding Modes and Thermodynamic Signatures of Closely Related Aldose Reductase Inhibitors, *J. Mol. Biol. 406*, 700-712.

(42) Steuber, H., Czodrowski, P., Sotriffer, C. A. and Klebe, G. (2007) Tracing Changes in Protonation: A Prerequisite to Factorize Thermodynamic Data of Inhibitor Binding to Aldose Reductase, *J. Mol. Biol. 373*, 1305-1320.

(43) Van Zandt, M. C., Sibley, E. O., McCann, E. E., Combs, K. J., Flam, B., Sawicki, D. R., Sabetta, A., Carrington, A., Sredy, J., Howard, E., Mitschler, A. and Podjarny, A. D. (2004) Design and Synthesis of Highly Potent and Selective (2-Arylcarbamoyl-phenoxy)-acetic Acid Inhibitors of Aldose Reductase for Treatment of Chronic Diabetic Complications, *Bioorg. Med. Chem. 12*, 5661-5675.

(44) Steuber, H. (2011) An Old NSAID Revisited: Crystal Structure of Aldose Reductase in Complex with Sulindac at 1.0 Å Supports a Novel Mechanism for Its Anticancer and Antiproliferative Effects, *ChemMedChem 6*, 2155-2157.

(45) Maccari, R. and Ottana, R. (2015) Targeting Aldose Reductase for the Treatment of Diabetes Complications and Inflammatory Diseases: New Insights and Future Directions, *J. Med. Chem. 58*, 2047-2067.

(46) Krishnamurthy, V. M., Kaufman, G. K., Urbach, A. R., Gitlin, I., Gudiksen, K. L., Weibel, D. B. and Whitesides, G. M. (2008) Carbonic Anhydrase as a Model for Biophysical and Physical-Organic Studies of Proteins and Protein−Ligand Binding, *Chem. Rev. 108*, 946-1051.

(47) Supuran, C. T. and Scozzafava, A. (2007) Carbonic Anhydrases as Targets for Medicinal Chemistry, *Bioorg. Med. Chem. 15*, 4336-4350.

(48) Chegwidden, W. R., Carter, N. D.; Edwards, Y. H., (Ed.) (2000) *Carbonic Anhydrases- New Horizons*, Vol. 90, Birkhäuser Verlag, Basel, Switzerland.

(49) Lindskog, S. (1997) Structure and Mechanism of Carbonic Anhydrase, *Pharmacol. Ther. 74*, 1-20.

(50) Liljas, A., Kannan, K. K., Bergsten, P. C., Waara, I., Fridborg, K., Strandberg, B., Carlbom, U., Jarup, L., Lovgren, S. and Petef, M. (1972) Crystal Structure of Human Carbonic Anhydrase C, *Nature (London), New Biol. 235*, 131-137.

(51) Tu, C. K., Silverman, D. N., Forsman, C., Jonsson, B. H. and Lindskog, S. (1989) Role of Histidine 64 in the Catalytic Mechanism of Human Carbonic Anhydrase II Studied with a Site-Specific Mutant, *Biochemistry 28*, 7913-7918.

(52) Briganti, F., Mangani, S., Orioli, P., Scozzafava, A., Vernaglione, G. and Supuran, C. T. (1997) Carbonic Anhydrase Activators: X-Ray Crystallographic and Spectroscopic Investigations for the Interaction of Isozymes I and II with Histamine, *Biochemistry 36*, 10384-10392.

(53) Avvaru, B. S., Kim, C. U., Sippel, K. H., Gruner, S. M., Agbandje-McKenna, M., Silverman, D. N. and McKenna, R. (2010) A Short, Strong Hydrogen Bond in the Active Site of Human Carbonic Anhydrase II, *Biochemistry 49*, 249-251.

(54) Supuran, C. T. and Scozzafava, A. (2000) Carbonic Anhydrase Inhibitors and Their Therapeutic Potential, *Expert Opin. Ther. Pat. 10*, 575-600.

(55) Kinsey, V. E. and Reddy, D. V. (1959) Turnover of Total Carbon Dioxide in the Aqueous Humors and the Effect Thereon of Acetazolamide, *AMA Arch. Ophthalmol. 62*, 78-83.

(56) Sugrue, M. F. (2000) Pharmacological and Ocular Hypotensive Properties of Topical Carbonic Anhydrase Inhibitors, *Prog. Retinal Eye Res. 19*, 87-112.

(57) Porcelli, M. J. and Gugelchuk, G. M. (1995) A Trek to the Top: A Review of Acute Mountain Sickness, *J. Am. Osteopath. Assoc. 95*, 718-720.

(58) Mecinovic, J., Snyder, P. W., Mirica, K. A., Bai, S., Mack, E. T., Kwant, R. L., Moustakas, D. T., Heroux, A. and Whitesides, G. M. (2011) Fluoroalkyl and Alkyl Chains Have Similar Hydrophobicities in Binding to the "Hydrophobic Wall" of Carbonic Anhydrase, *JACS 133*, 14017-14026.

(59) Whitesides, G. M. and Krishnamurthy, V. M. (2005) Designing Ligands to Bind Proteins, *Q. Rev. Biophys. 38*, 385-395.

(60) Myszka, D. G. (2004) Analysis of Small-Molecule Interactions Using Biacore S51 Technology, *Anal. Biochem. 329*, 316-323.

(61) Navratilova, I., Papalia, G. A., Rich, R. L., Bedinger, D., Brophy, S., Condon, B., Deng, T., Emerick, A. W., Guan, H. W., Hayden, T., Heutmekers, T., Hoorelbeke, B., McCroskey, M. C., Murphy, M. M., Nakagawa, T., Parmeggiani, F., Qin, X., Rebe, S., Tomasevic, N., Tsang, T., Waddell, M. B., Zhang, F. F., Leavitt, S. and Myszka, D. G. (2007) Thermodynamic Benchmark Study Using Biacore Technology, *Anal. Biochem. 364*, 67-77.

(62) Cheng, T., Li, X., Li, Y., Liu, Z. and Wang, R. (2009) Comparative Assessment of Scoring Functions on a Diverse Test Set, *J. Chem. Inf. Model. 49*, 1079-1093.

(63) Schneider, N., Lange, G., Hindle, S., Klein, R. and Rarey, M. (2013) A Consistent Description of Hydrogen Bond and Dehydration Energies in Protein-Ligand Complexes: Methods Behind the Hyde Scoring Function, *J. Comput. Aided Mol. Des. 27*, 15-29.

(64) Ramirez, M. A. and Borja, N. L. (2008) Epalrestat: An Aldose Reductase Inhibitor for the Treatment of Diabetic Neuropathy, *Pharmacotherapy 28*, 646-655.

(65) Kraemer, O., Hazemann, I., Podjarny, A. D. and Klebe, G. (2004) Virtual Screening for Inhibitors of Human Aldose Reductase, *Proteins: Struct., Funct., Bioinf. 55*, 814-823.

(66) King, R. W. and Burgen, A. S. V. (1976) Kinetic Aspects of Structure-Activity Relations: The Binding of Sulphonamides by Carbonic Anhydrase, *Proc. R. Soc. B 193*, 107-125.

(67) Taylor, P. W., King, R. W. and Burgen, A. S. V. (1970) Kinetics of Complex Formation between Human Carbonic Anhydrases and Aromatic Sulfonamides, *Biochemistry 9*, 2638-2645.

(68) Yabe-Nishimura, C. (1998) Aldose Reductase in Glucose Toxicity: A Potential Target for the Prevention of Diabetic Complications, *Pharmacol. Rev. 50*, 21-33.

(69) Srivastava, S. K., Ramana, K. V. and Bhatnagar, A. (2005) Role of Aldose Reductase and Oxidative Damage in Diabetes and the Consequent Potential for Therapeutic Options, *Endocr. Rev. 26*, 380-392.

(70) Ramana, K. V. (2011) Aldose Reductase: New Insights for an Old Enzyme, *Biomol. Concepts 2*, 103-114.

(71) Ramana, K. V. and Srivastava, S. K. (2010) Aldose Reductase: A Novel Therapeutic Target for Inflammatory Pathologies, *Int. J. Biochem. Cell Biol. 42*, 17-20.

(72) Hotta, N. (1995) New Approaches for Treatment in Diabetes: Aldose Reductase Inhibitors, *Biomed. Pharmacother. 5*, 244-250.

(73) Costantino, L., Rastelli, G., Gamberini, M. C. and Barlocco, D. (2000) Pharmacological Approaches to the Treatment of Diabetic Complications, *Expert Opin. Ther. Pat. 10*, 1245-1262.

(74) Miyamoto, S. (2002) Recent Advances in Aldose Reductase Inhibitors: Potential Agents for the Treatment of Diabetic Complications, *Expert Opin. Ther. Pat. 12*, 621-631.

(75) Chatzopoulou, M., Pegklidou, K., Papastavrou, N. and Demopoulos, V. J. (2013) Development of Aldose Reductase Inhibitors for the Treatment of Inflammatory Disorders, *Expert Opin. Drug Discovery 8*, 1365-1380.

(76) Alexiou, P., Pegklidou, K., Chatzopoulou, M., Nicolaou, I. and Demopoulos, V. J. (2009) Aldose Reductase Enzyme and Its Implication to Major Health Problems of the 21st Century, *Curr. Med. Chem. 16*, 734-752.

(77) Chatzopoulou, M., Alexiou, P., Kotsampasakou, E. and Demopoulos, V. J. (2012) Novel Aldose Reductase Inhibitors: A Patent Survey (2006-Present), *Expert Opin. Ther. Pat. 11*, 1303-1323.

(78) Del-Corso, A., Balestri, F., Di Bugno, E., Moschini, R., Cappiello, M., Sartini, S., La-Motta, C., Da-Settimo, F. and Mura, U. (2013) A New Approach to Control the Enigmatic Activity of Aldose Reductase, *PLoS One 8*, e74076.

(79) Cappiello, M., Moschini, R., Balestri, F., Mura, U. and Del-Corso, A. (2014) Basic Models for Differential Inhibition of Enzymes, *Biochem. Biophys. Res. Commun. 445*, 556-560.

(80) Van Zandt, M. C., Doan, B., Sawicki, D. R., Sredy, J. and Podjarny, A. D. (2009) Discovery of [3-(4,5,7-Trifluoro-benzothiazol-2-ylmethyl)-pyrrolo[2,3-B]pyridin-1-yl]acetic Acids as Highly Potent and Selective Inhibitors of Aldose Reductase for Treatment of Chronic Diabetic Complications, *Bioorg. Med. Chem. Lett. 19*, 2006-2008.

(81) Van Zandt, M. C., Jones, M. L., Gunn, D. E., Geraci, L. S., Jones, J. H., Sawicki, D. R., Sredy, J., Jacot, J. L., DiCioccio, A. T., Petrova, T., Mitschler, A. and Podjarny, A. D. (2005) Discovery of 3-[(4,5,7-Trifluorobenzothiazol-2-yl)methyl]indole-N-acetic Acid (Lidorestat) and Congeners as Highly Potent and Selective Inhibitors of Aldose Reductase for Treatment of Chronic Diabetic Complications, *J. Med. Chem. 48*, 3141-3152.

(82) Toyooka, N., Kato, A. and Adachi, I., Preparation of Fused Tricyclic Compounds as Aldose Reductase Inhibitors (2009), National University Corporation of Toyama, Japan, WO2009078423A1

(83) Luker, T., Bonnert, R., Brough, S., Cook, A. R., Dickinson, M. R., Dougall, I., Logan, C., Mohammed, R. T., Paine, S., Sanganee, H. J., Sargent, C., Schmidt, J. A., Teague, S. and Thom, S. (2011) Substituted Indole-1-acetic Acids as Potent and Selective CRTH2 Antagonists-Discovery of AZD1981, *Bioorg. Med. Chem. Lett. 21*, 6288-6292.

(84) Chung, S. and LaMendola, J. (1989) Cloning and Sequence Determination of Human Placental Aldose Reductase Gene, *J. Biol. Chem. 264*, 14775-14777.

(85) Gui, T., Tanimoto, T., Kokai, Y. and Nishimura, C. (1995) Presence of a Closely Related Subgroup in the Aldo-Ketoreductase Family of the Mouse, *Eur. J. Biochem. 227*, 448-453.

(86) *Marvin* (2014), 14.7.21.0, ChemAxon, Budapest, Hungary.

(87) Fukada, H. and Takahashi, K. (1998) Enthalpy and Heat Capacity Changes for the Proton Dissociation of Various Buffer Components in 0.1 M Potassium Chloride, *Proteins: Struct., Funct., Genet. 33*, 159-166.

(88) Dullweber, F., Stubbs, M. T., Musil, D., Sturzebecher, J. and Klebe, G. (2001) Factorising Ligand Affinity: A Combined Thermodynamic and Crystallographic Study of Trypsin and Thrombin Inhibition, *J. Mol. Biol. 313*, 593-614.

(89) *ACD/Chemsketch* (2014), 2013-11-07, Advanced Chemistry Development, Inc., Toronto, Ontario, Canada.

(90) Hopkins, A. L., Groom, C. R. and Alex, A. (2004) Ligand Efficiency: A Useful Metric for Lead Selection, *Drug Discovery Today 9*, 430-431.

(91) Keseru, G. M. and Makara, G. M. (2009) The Influence of Lead Discovery Strategies on the Properties Drug Candidates, *Nat. Rev. Drug Discov 8*, 203-212.

(92) Perola, E. (2010) An Analysis of the Binding Efficiencies of Drugs and Their Leads in Successful Drug Discovery Programs, *J. Med. Chem. 53*, 2986-2997.

(93) Terashima, H., Hama, K., Yamamoto, R., Tsuboshima, M., Kikkawa, R., Hatanaka, I. and Shigeta, Y. (1984) Effects of a New Aldose Reductase Inhibitor on Various Tissues in Vitro, *J. Pharmacol. Exp. Ther. 229*, 226-230.

(94) Chambers, C. C., Cramer, C. J. and Truhlar, D. G. (1996) Model for Aqueous Solvation Based on Class IV Atomic Charges and First Solvation Shell Effects, *J. Phys. Chem. 100*, 16385-16398.

(95) Hayman, S. and Kinoshita, J. H. (1965) Isolation and Properties of Lens Aldose Reductase, *J. Biol. Chem. 240*, 877-882.

(96) Costantino, L., Rastelli, G., Gamberini, M. C., Vinson, J. A., Bose, P., Iannone, A., Staffieri, M., Antolini, L., Del Corso, A., Mura, U. and Albasini, A. (1999) 1-Benzopyran-4-one Antioxidants as Aldose Reductase Inhibitors, *J. Med. Chem. 42*, 1881-1893.

(97) Stefek, M., Snirc, V., Djoubissie, P. O., Majekova, M., Demopoulos, V., Rackova, L., Bezakova, Z., Karasu, C., Carbone, V. and El-Kabbani, O. (2008) Carboxymethylated Pyridoindole Antioxidants as Aldose Reductase Inhibitors: Synthesis, Activity, Partitioning, and Molecular Modeling, *Bioorg. Med. Chem. 16*, 4908-4920.

(98) Stefek, M., Tsantili-Kakoulidou, A., Milackova, I., Juskova, M., Snirc, V. and Triantos, N. (2011) (2-Benzyl-2,3,4,5-tetrahydro-1H-pyrido[4,3-B]indol-8-yl)-acetic Acid: An Aldose Reductase Inhibitor and Antioxidant of Zwitterionic Nature, *Bioorg. Med. Chem. 19*, 7181-7185.

(99) Mylari, B. L., Armento, S. J., Beebe, D. A., Conn, E. L., Coutcher, J. B., Dina, M. S., O'Gorman, M. T., Linhares, M. C., Martin, W. H., Oates, P. J., Tess, D. A., Withbroe, G. J. and Zembrowski, W. J. (2003) A Highly Selective, Non-Hydantoin, Non-Carboxylic Acid Inhibitor of Aldose Reductase with Potent Oral Activity in Diabetic Rat Models: 6-(5-Chloro-3-methylbenzofuran-2-sulfonyl)-2-H-pyridazin-3-one, *J. Med. Chem. 46*, 2283-2286.

(100) Sangster, J. *In Octanol–Water Partition Coefficients: Fundamentals and Physical Chemistry* (1997), John Willey & Sons, England.

(101) Cousido-Siah, A., Ruiz, F. X., Mitschler, A., Porte, S., de Lera, A. R., Martin, M. J., Manzanaro, S., de la Fuente, J. A., Terwesten, F., Betz, M., Klebe, G., Farres, J., Pares, X. and Podjarny, A. (2014) Identification of a Novel Polyfluorinated Compound as a Lead to Inhibit the Human Enzymes Aldose Reductase and AKR1B10: Structure Determination of Both Ternary Complexes and Implications for Drug Design, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 70*, 889-903.

(102) Fukumoto, S., Yamauchi, N., Moriguchi, H., Hippo, Y., Watanabe, A., Shibahara, J., Taniguchi, H., Ishikawa, S., Ito, H., Yamamoto, S., Iwanari, H., Hironaka, M., Ishikawa, Y., Niki, T., Sohara, Y., Kodama, T., Nishimura, M., Fukayama, M., Dosaka-Akita, H. and Aburatani, H. (2005) Overexpression of the Aldo-Keto Reductase Family Protein AKR1B10 Is Highly Correlated with Smokers' Non-Small Cell Lung Carcinomas, *Clin. Cancer Res. 11*, 1776-1785.

(103) Ruiz, F. X., Gallego, O., Ardèvol, A., Moro, A., Domínguez, M., Álvarez, S., Álvarez, R., de Lera, A. R., Rovira, C., Fita, I., Parés, X. and Farrés, J. (2009) Aldo-Keto Reductases from the AKR1B Subfamily: Retinoid Specificity and Control of Cellular Retinoic Acid Levels, *Chem.-Biol. Interact 178*, 171-177.

(104) Wang, C., Yan, R., Luo, D., Watabe, K., Liao, D. F. and Cao, D. (2009) Aldo-Keto Reductase Family 1 Member B10 Promotes Cell Survival by Regulating Lipid Synthesis and Eliminating Carbonyls, *J. Biol. Chem. 284*, 26742-26748.

(105) Zeindl-Eberhart, E., Haraida, S., Liebmann, S., Jungblut, P. R., Lamer, S., Mayer, D., Jager, G., Chung, S. and Rabes, H. M. (2004) Detection and Identification of Tumor-Associated Protein Variants in Human Hepatocellular Carcinomas, *Hepatology 39*, 540-549.

(106) Barski, O. A., Tipparaju, S. M. and Bhatnagar, A. (2008) The Aldo-Keto Reductase Superfamily and Its Role in Drug Metabolism and Detoxification, *Drug Metab. Rev. 40*, 553-624.

(107) Gallego, O., Belyaeva, O. V., Porté, S., Ruiz, F. X., Stetsenko, A. V., Shabrova, E. V., Kostereva, N. V., Farrés, J., Parés, X. and Kedishvili, N. Y. (2006) Comparative Functional Analysis of Human Medium-Chain Dehydrogenases, Short-Chain Dehydrogenases/Reductases and Aldo-Keto Reductases with Retinoids, *Biochem. J. 399*, 101-109.

(108) Gallego, O., Ruiz, F. X., Ardèvol, A., Domínguez, M., Álvarez, R., de Lera, A. R., Rovira, C., Farrés, J., Fita, I. and Parés, X. (2007) Structural Basis for the High All-Trans-Retinaldehyde Reductase Activity of the Tumor Marker AKR1B10, *Proc. Natl. Acad. Sci. U. S. A. 104*, 20764-20769.

(109) Penning, T. M. (2005) AKR1B10: A New Diagnostic Marker of Non-Small Cell Lung Carcinoma in Smokers, *Clin. Cancer Res. 11*, 1687-1690.

(110) Quinn, A. M., Harvey, R. G. and Penning, T. M. (2008) Oxidation of PAH Trans-Dihydrodiols by Human Aldo-Keto Reductase AKR1B10, *Chem. Res. Toxicol. 21*, 2207-2215.

(111) Ruiz, F. X., Porté, S., Parés, X. and Farrés, J. (2012) Biological Role of Aldo-Keto Reductases in Retinoic Acid Biosynthesis and Signaling, *Front. Pharmacol. 3*, 1-13.

(112) Ma, J., Yan, R., Zu, X., Cheng, J. M., Rao, K., Liao, D. F. and Cao, D. (2008) Aldo-Keto Reductase Family 1 B10 Affects Fatty Acid Synthesis by Regulating the Stability of Acetyl-CoA Carboxylase-Alpha in Breast Cancer Cells, *J. Biol. Chem. 283*, 3418-3423.

(113) de la Fuente, J. A. and Manzanaro, S. (2003) Aldose Reductase Inhibitors from Natural Sources, *Nat. Prod. Rep. 20*, 243-251.

(114) de la Fuente, J. A., Manzanaro, S., Martín, M. J., de Quesada, T. G., Reymundo, I., Luengo, S. M. and Gago, F. (2003) Synthesis, Activity, and Molecular Modeling Studies of Novel Human Aldose Reductase Inhibitors Based on a Marine Natural Product, *J. Med. Chem. 46*, 5208-5221.

(115) Manzanaro, S., Salva, J. and de la Fuente, J. A. (2006) Phenolic Marine Natural Products as Aldose Reductase Inhibitors, *J. Nat. Prod. 69*, 1485-1487.

(116) Kotani, T., Nagaki, Y., Ishii, A., Konishi, Y., Yago, H., Suehiro, S., Okukado, N. and Okamoto, K. (1997) Highly Selective Aldose Reductase Inhibitors. 3. Structural Diversity of 3-(Arylmethyl)-2,4,5-trioxoimidazolidine-1-acetic Acids, *J. Med. Chem. 40*, 684-694.

(117) Bohren, K. M. and Grimshaw, C. E. (2000) The Sorbinil Trap: A Predicted Dead-End Complex Confirms the Mechanism of Aldose Reductase Inhibition, *Biochemistry 39*, 9967-9974.

(118) El-Kabbani, O. and Podjarny, A. (2007) Selectivity Determinants of the Aldose and Aldehyde Reductase Inhibitor-Binding Sites, *Cell. Mol. Life. Sci. 64*, 1970-1978.

(119) Blakeley, M. P., Ruiz, F., Cachau, R., Hazemann, I., Meilleur, F., Mitschler, A., Ginell, S., Afonine, P., Ventura, O. N., Cousido-Siah, A., Haertlein, M., Joachimiak, A., Myles, D. and Podjarny, A. (2008) Quantum Model of Catalysis Based on a Mobile Proton Revealed by Subatomic X-Ray and Neutron Diffraction Studies of H-Aldose Reductase, *Proc. Natl. Acad. Sci. U. S. A. 105*, 1844-1848.

(120) Bissantz, C., Kuhn, B. and Stahl, M. (2010) A Medicinal Chemist's Guide to Molecular Interactions, *J. Med. Chem. 53*, 5061-5084.

(121) Ruiz, F., Hazemann, I., Mitschler, A., Joachimiak, A., Schneider, T., Karplus, M. and Podjarny, A. (2004) The Crystallographic Structure of the Aldose Reductase-IDD552 Complex Shows Direct Proton Donation from Tyrosine 48, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 60*, 1347-1354.

(122) Borhani, D. W., Harter, T. M. and Petrash, J. M. (1992) The Crystal Structure of the Aldose Reductase.NADPH Binary Complex, *J. Biol. Chem. 267*, 24841-24847.

(123) Zhang, L., Zhang, H., Zhao, Y., Li, Z., Chen, S., Zhai, J., Chen, Y., Xie, W., Wang, Z., Li, Q., Zheng, X. and Hu, X. (2013) Inhibitor Selectivity between Aldo-Keto Reductase Superfamily Members AKR1B10 and AKR1B1: Role of Trp112 (Trp111), *FEBS Lett. 587*, 3681–3686.

(124) Zhang, L., Zhang, H., Zheng, X., Zhao, Y., Chen, S., Chen, Y., Zhang, R., Li, Q. and Hu, X. (2014) Structural Basis for the Inhibition of AKR1B10 by Caffeic Acid Phenethyl Ester (CAPE), *ChemMedChem*, 706-709.

(125) Cousido-Siah, A., Ruiz, F. X., Crespo, I., Porté, S., Mitschler, A., Parés, X., Podjarny, A. and Farrés, J. (2014) Structural Analysis of Sulindac as an Inhibitor of Aldose Reductase and AKR1B10, *Chem.- Biol. Interact. 234*, 290-296.

(126) Derewenda, Z. S., Lee, L. and Derewenda, U. (1995) The Occurence of C-H ⋯ O Hydrogen Bonds in Proteins, *J. Mol. Biol. 252*, 248-262.

(127) Stefek, M., Soltesova Prnova, M., Majekova, M., Rechlin, C., Heine, A. and Klebe, G. (2015) Identification of Novel Aldose Reductase Inhibitors Based on Carboxymethylated Mercaptotriazinoindole Scaffold, *J. Med. Chem. 58*, 2649-2657.

(128) Maccari, R. and Ottana, R. (2015) Targeting Aldose Reductase for the Treatment of Diabetes Complications and Inflammatory Diseases: New Insights and Future Directions, *J. Med. Chem. 58*, 2047-2067.

(129) Ma, J., Luo, D. X., Huang, C., Shen, Y., Bu, Y., Markwell, S., Gao, J., Liu, J., Zu, X., Cao, Z., Gao, Z., Lu, F., Liao, D. F. and Cao, D. (2012) AKR1B10 Overexpression in Breast Cancer: Association with Tumor Size, Lymph Node Metastasis and Patient Survival and Its Potential as a Novel Serum Marker, *Int. J. Cancer 131*, E862-E871.

(130) Satow, R., Shitashige, M., Kanai, Y., Takeshita, F., Ojima, H., Jigami, T., Honda, K., Kosuge, T., Ochiya, T., Hirohashi, S. and Yamada, T. (2010) Combined Functional Genome Survey of Therapeutic Targets for Hepatocellular Carcinoma, *Clin. Cancer Res. 16*, 2518-2528.

(131) Tammali, R., Ramana, K. V., Singhal, S. S., Awasthi, S. and Srivastava, S. K. (2006) Aldose Reductase Regulates Growth Factor-Induced Cyclooxygenase-2 Expression and Prostaglandin E2 Production in Human Colon Cancer Cells, *Cancer Res. 66*, 9705-9713.

(132) Tammali, R., Reddy, A. B., Saxena, A., Rychahou, P. G., Evers, B. M., Qiu, S., Awasthi, S., Ramana, K. V. and Srivastava, S. K. (2011) Inhibition of Aldose Reductase Prevents Colon Cancer Metastasis, *Carcinogenesis 32*, 1259-1267.

(133) Zhang, W., Li, H., Yang, Y., Liao, J. and Yang, G.-Y. (2014) Knockdown or Inhibition of Aldo-Keto Reductase 1B10 Inhibits Pancreatic Carcinoma Growth Via Modulating Kras-E-Cadherin Pathway, *Cancer Lett. 355*, 273-280.

(134) Lamour, V., Barth, P., Rogniaux, H., Poterszman, A., Howard, E., Mitschler, A., Van Dorsselaer, A., Podjarny, A. and Moras, D. (1999) Production of Crystals of Human Aldose Reductase with Very High Resolution Diffraction, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 55*, 721-723.

(135) Barski, O. A., Gabbay, K. H., Grimshaw, C. E. and Bohren, K. M. (1995) Mechanism of Human Aldehyde Reductase: Characterization of the Active Site Pocket, *Biochemistry 34*, 11264-11275.

(136) Cousido-Siah, A., Petrova, T., Hazemann, I., Mitschler, A., Ruiz, F. X., Howard, E., Ginell, S., Atmanene, C., Van Dorsselaer, A., Sanglier-Cianferani, S., Joachimiak, A. and Podjarny, A. (2012) Crystal Packing Modifies Ligand Binding Affinity: The Case of Aldose Reductase, *Proteins 80*, 2552-2561.

(137) Otwinowski, Z. and Minor, W. (1997) Processing of X-Ray Diffraction Data Collected in Oscillation Mode, In *Methods Enzymol.* (Charles W. Carter, Jr., Ed.), 307-326, Academic Press.

(138) McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. and Read, R. J. (2007) Phaser Crystallographic Software, *J. Appl. Crystallogr. 40*, 658-674.

(139) Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K. (2010) Features and Development of Coot, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 486-501.

(140) Murshudov, G. N., Skubak, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. and Vagin, A. A. (2011) REFMAC5 for the Refinement of Macromolecular Crystal Structures, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 67*, 355-367.

(141) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. and Zwart, P. H. (2010) Phenix: A Comprehensive Python-Based System for Macromolecular Structure Solution, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 213-221.

(142) Moriarty, N. W., Grosse-Kunstleve, R. W. and Adams, P. D. (2009) Electronic Ligand Builder and Optimization Workbench (Elbow): A Tool for Ligand Coordinate and Restraint Generation, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 65*, 1074-1080.

(143) Eisenmann, M., Steuber, H., Zentgraf, M., Altenkamper, M., Ortmann, R., Perruchon, J., Klebe, G. and Schlitzer, M. (2009) Structure-Based Optimization of Aldose Reductase Inhibitors Originating from Virtual Screening, *ChemMedChem 4*, 809-819.

(144) Koch, C. (2011) Towards Improved Aldose Reductase Inhibitors- Structural and Thermodynamic Investigation of Mutant and Wild Type Aldose Reductase Inhibitor Complexes, Philipps Universität Marburg.

(145) Koch, C., Heine, A. and Klebe, G. (2011) Ligand-Induced Fit Affects Binding Modes and Provokes Changes in Crystal Packing of Aldose Reductase, *Biochim. Biophys. Acta 1810*, 879-887.

(146) Sturtevant, J. M. (1977) Heat Capacity and Entropy Changes in Processes Involving Proteins, *PNAS 74*, 2236-2240.

(147) Tanford, C. (1978) The Hydrophobic Effect and the Organization of Living Matter, *Science 200*, 1012-1018.

(148) Snyder, P., Lockett, M., Moustakas, D. and Whitesides, G. (2014) Is It the Shape of the Cavity, or the Shape of the Water in the Cavity?, *Eur. Phys. J.: Spec. Top. 223*, 853-891.

(149) http://www.molinspiration.com. Molinspiration Property Engine, V2014.11, *last access: 06.08.2015*.

(150) Nishikawa, K., Fukuda, H., Abe, M., Nakanishi, K., Taniguchi, T., Nomura, T., Yamaguchi, C., Hiradate, S., Fujii, Y., Okuda, K. and Shindo, M. (2013) Substituent Effects of Cis-Cinnamic Acid Analogues as Plant Growh Inhibitors, *Phytochemistry 96*, 132-147.

(151) El-Kabbani, O., Darmanin, C., Schneider, T. R., Hazemann, I., Ruiz, F., Oka, M., Joachimiak, A., Schulze-Briese, C., Tomizaki, T., Mitschler, A. and Podjarny, A. (2004) Ultrahigh Resolution Drug Design. II. Atomic Resolution Structures of Human Aldose Reductase Holoenzyme Complexed with Fidarestat and Minalrestat: Implications for the Binding of Cyclic Imide Inhibitors, *Proteins 55*, 805-813.

(152) McLachlan, A. (1982) Rapid Comparison of Protein Structures, *Acta Crystallogr., Sect. A 38*, 871-873.

(153) Martin, A. C. R. and Porter, C. T. *Profit*, V3.1.

(154) Wlodawer, A., Minor, W., Dauter, Z. and Jaskolski, M. (2008) Protein Crystallography for Non-Crystallographers, or How to Get the Best (but Not More) from Published Macromolecular Structures, *FEBS J. 275*, 1-21.

(155) Baum, B., Mohamed, M., Zayed, M., Gerlach, C., Heine, A., Hangauer, D. and Klebe, G. (2009) More Than a Simple Lipophilic Contact: A Detailed Thermodynamic Analysis of Nonbasic Residues in the S1 Pocket of Thrombin, *J. Mol. Biol. 390*, 56-69.

(156) Biela, A., Khayat, M., Tan, H., Kong, J., Heine, A., Hangauer, D. and Klebe, G. (2012) Impact of Ligand and Protein Desolvation on Ligand Binding to the S1 Pocket of Thrombin, *J. Mol. Biol. 418*, 350-366.

(157) Rühmann, E., Betz, M., Fricke, M., Heine, A., Schäfer, M. and Klebe, G. (2015) Thermodynamic Signatures of Fragment Binding: Validation of Direct Versus Displacement ITC Titrations, *Biochim. Biophys. Acta 1850*, 647-656.

(158) Sigurskjold, B. W. (2000) Exact Analysis of Competition Ligand Binding by Displacement Isothermal Titration Calorimetry, *Anal. Biochem. 277*, 260-266.

(159) Zhang, Y. L. and Zhang, Z. Y. (1998) Low-Affinity Binding Determined by Titration Calorimetry Using a High-Affinity Coupling Ligand: A Thermodynamic Study of Ligand Binding to Protein Tyrosine Phosphatase 1B, *Anal. Biochem. 261*, 139-148.

(160) Keller, S., Vargas, C., Zhao, H., Piszczek, G., Brautigam, C. A. and Schuck, P. (2012) High-Precision Isothermal Titration Calorimetry with Automated Peak-Shape Analysis, *Anal. Chem. 84*, 5066-5073.

(161) Zhao, H., Piszczek, G. and Schuck, P. (2015) Sedphat - a Platform for Global ITC Analysis and Global Multi-Method Analysis of Molecular Interactions, *Methods 76*, 137-148.

(162) Supuran, C. T. (2008) Carbonic Anhydrases: Novel Therapeutic Applications for Inhibitors and Activators, *Nat. Rev. Drug Discovery 7*, 168-181.

(163) Gao, J., Qiao, S. and Whitesides, G. M. (1995) Increasing Binding Constants of Ligands to Carbonic Anhydrase by Using "Greasy Tails", *J. Med. Chem. 38*, 2292-2301.

(164) Snyder, P. W., Mecinovic, J., Moustakas, D. T., Thomas, S. W., 3rd, Harder, M., Mack, E. T., Lockett, M. R., Heroux, A., Sherman, W. and Whitesides, G. M. (2011) Mechanism of the Hydrophobic Effect in the Biomolecular Recognition of Arylsulfonamides by Carbonic Anhydrase, *Proc. Natl. Acad. Sci. U. S. A. 108*, 17889-17894.

(165) Kockar, F., Maresca, A., Aydın, M., Işık, S., Turkoglu, S., Sinan, S., Arslan, O., Güler, Ö. Ö., Turan, Y. and Supuran, C. T. (2010) Mutation of Phe91 to Asn in Human Carbonic Anhydrase I Unexpectedly Enhanced Both Catalytic Activity and Affinity for Sulfonamide Inhibitors, *Bioorg. Med. Chem. 18*, 5498-5503.

(166) Schmid, M., Nogueira, E. S., Monnard, F. W., Ward, T. R. and Meuwly, M. (2012) Arylsulfonamides as Inhibitors for Carbonic Anhydrase: Prediction & Validation, *Chem. Sci. 3*, 690-700.

(167) Turkoglu, S., Maresca, A., Alper, M., Kockar, F., Işık, S., Sinan, S., Ozensoy, O., Arslan, O. and Supuran, C. T. (2012) Mutation of Active Site Residues Asn67 to Ile, Gln92 to Val and Leu204 to Ser in Human Carbonic Anhydrase II: Influences on the Catalytic Activity and Affinity for Inhibitors, *Bioorg. Med. Chem. 20*, 2208-2213.

(168) Kanamori, K. and Roberts, J. D. (1983) Nitrogen-15 Nuclear Magnetic Resonance Study of Benzenesulfonamide and Cyanate Binding to Carbonic Anhydrase, *Biochemistry 22*, 2658-2664.

(169) Fisher, S. Z., Aggarwal, M., Kovalevsky, A. Y., Silverman, D. N. and McKenna, R. (2012) Neutron Diffraction of Acetazolamide-Bound Human Carbonic Anhydrase II Reveals Atomic Details of Drug Binding, *J. Am. Chem. Soc. 134*, 14726-14729.

(170) Kiefer, L. L., Paterno, S. A. and Fierke, C. A. (1995) Hydrogen Bond Network in the Metal Binding Site of Carbonic Anhydrase Enhances Zinc Affinity and Catalytic Efficiency, *J. Am. Chem. Soc. 117*, 6831-6837.

(171) Taylor, P. W., King, R. W. and Burgen, A. S. (1970) Influence of pH on the Kinetics of Complex Formation between Aromatic Sulfonamides and Human Carbonic Anhydrase, *Biochemistry 9*, 3894-3902.

(172) Eriksson, A. E., Jones, T. A. and Liljas, A. (1988) Refined Structure of Human Carbonic Anhydrase II at 2.0 Å Resolution, *Proteins 4*, 274-282.

(173) Boriack, P. A., Christianson, D. W., Kingery-Wood, J. and Whitesides, G. M. (1995) Secondary Interactions Significantly Removed from the Sulfonamide Binding Pocket of Carbonic Anhydrase II Influence Inhibitor Binding Constants, *J. Med. Chem. 38*, 2286-2291.

(174) Scott, A. D., Phillips, C., Alex, A., Flocco, M., Bent, A., Randall, A., O'Brien, R., Damian, L. and Jones, L. H. (2009) Thermodynamic Optimisation in Drug Discovery: A Case Study Using Carbonic Anhydrase Inhibitors, *ChemMedChem 4*, 1985-1989.

(175) Martin, D. P., Hann, Z. S. and Cohen, S. M. (2013) Metalloprotein–Inhibitor Binding: Human Carbonic Anhydrase II as a Model for Probing Metal–Ligand Interactions in a Metalloprotein Active Site, *Inorg. Chem. 52*, 12207-12215.

(176) Aggarwal, M., Boone, C. D., Kondeti, B., Tu, C., Silverman, D. N. and McKenna, R. (2013) Effects of Cryoprotectants on the Structure and Thermostability of the Human Carbonic Anhydrase II-Acetazolamide Complex, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 69*, 860-865.

(177) Lee, B. and Richards, F. M. (1971) The Interpretation of Protein Structures: Estimation of Static Accessibility, *J. Mol. Biol. 55*, 379-400.

(178) van Holde, K. E. (2002) A Hypothesis Concerning Diffusion-Limited Protein–Ligand Interactions, *Biophys. Chem. 101–102*, 249-254.

(179) Reynolds, J. A., Gilbert, D. B. and Tanford, C. (1974) Empirical Correlation between Hydrophobic Free Energy and Aqueous Cavity Surface Area, *Proc. Natl. Acad. Sci. U. S. A. 71*, 2925-2927.

(180) Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A. and Simmerling, C. (2006) Comparison of Multiple Amber Force Fields and Development of Improved Protein Backbone Parameters, *Proteins 65*, 712-725.

(181) Peters, M. B., Yang, Y., Wang, B., Füsti-Molnár, L., Weaver, M. N. and Merz, K. M. (2010) Structural Survey of Zinc-Containing Proteins and Development of the Zinc Amber Force Field (ZAFF), *J. Chem. Theory Comput. 6*, 2935-2947.

(182) Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. and Case, D. A. (2004) Development and Testing of a General Amber Force Field, *J. Comput. Chem. 25*, 1157-1174.

(183) Bayly, C. I., Cieplak, P., Cornell, W. and Kollman, P. A. (1993) A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges: The RESP Model, *J. Phys. Chem. 97*, 10269-10280.

(184) *Macromodel* (2012), 9.9, Schrödinger, LLC, New York.

(185) Jorgensen, W. L. and Tirado-Rives, J. (1988) The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin, *J. Am. Chem. Soc. 110*, 1657-1666.

(186) Jorgensen, W. L., Maxwell, D. S. and Tirado-Rives, J. (1996) Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids, *J. Am. Chem. Soc. 118*, 11225-11236.

(187) Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M. R., Smith, J. C., Kasson, P. M., van der Spoel, D., Hess, B. and Lindahl, E. (2013) GROMACS 4.5: A High-Throughput and Highly Parallel Open Source Molecular Simulation Toolkit, *Bioinformatics 29*, 845-854.

(188) Parrinello, M. and Rahman, A. (1981) Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method, *J. Appl. Phys. 52*, 7182-7190.

(189) Bussi, G., Donadio, D. and Parrinello, M. (2007) Canonical Sampling through Velocity Rescaling, *J. Chem. Phys. 126*, 014101.

(190) Barducci, A., Bussi, G. and Parrinello, M. (2008) Well-Tempered Metadynamics: A Smoothly Converging and Tunable Free-Energy Method, *Phys. Rev. Lett. 100*, 020603.

(191) Bonomi, M., Branduardi, D., Bussi, G., Camilloni, C., Provasi, D., Raiteri, P., Donadio, D., Marinelli, F., Pietrucci, F., Broglia, R. A. and Parrinello, M. (2009) PLUMED: A Portable Plugin for Free-Energy Calculations with Molecular Dynamics, *Comput. Phys. Commun. 180*, 1961-1972.

(192) Daura, X., Gademann, K., Jaun, B., Seebach, D., van Gunsteren, W. F. and Mark, A. E. (1999) Peptide Folding: When Simulation Meets Experiment, *Angew. Chem., Int. Ed. 38*, 236-240.

(193) Rocchia, W., Alexov, E. and Honig, B. (2001) Extending the Applicability of the Nonlinear Poisson−Boltzmann Equation: Multiple Dielectric Constants and Multivalent Ions, *J. Phys. Chem. B 105*, 6507-6514.

(194) Hutter, J., Iannuzzi, M., Schiffmann, F. and VandeVondele, J. (2014) CP2K: Atomistic Simulations of Condensed Matter Systems, *WIREs Comput. Mol. Sci. 4*, 15-25.

(195) Goedecker, S., Teter, M. and Hutter, J. (1996) Separable Dual-Space Gaussian Pseudopotentials, *Phys. Rev. B 54*, 1703-1710.

(196) VandeVondele, J. and Hutter, J. (2007) Gaussian Basis Sets for Accurate Calculations on Molecular Systems in Gas and Condensed Phases, *J. Chem. Phys. 127*, 114105-114105.

(197) Hellriegel, C. and Rueck, A. (2015) *Sigma Aldrich Analytix 1*, 6.

(198) Pauli, G. F., Chen, S.-N., Simmler, C., Lankin, D. C., Gödecke, T., Jaki, B. U., Friesen, J. B., McAlpine, J. B. and Napolitano, J. G. (2014) Importance of Purity Evaluation and the Potential of Quantitative 1H NMR as a Purity Assay, *J. Med. Chem. 57*, 9220-9231.

(199) Steuber, H. (2007) Structural and Thermodynamic Characterization of Inhibitor Binding to Aldose Reductase: Insights into Binding Modes, Driving Forces, and Selectivity Determinants, Philipps-Universität Marburg.

(200) Krämer, O. (2003) Theoretische und Experimentelle Untersuchungen zur Inhibition des Enzyms Aldose Reduktase, Philipps-Universität Marburg.

(201) Mueller, U., Darowski, N., Fuchs, M. R., Forster, R., Hellmig, M., Paithankar, K. S., Puhringer, S., Steffien, M., Zocher, G. and Weiss, M. S. (2012) Facilities for Macromolecular Crystallography at the Helmholtz-Zentrum Berlin, *J. Synchrotron Radiat. 19*, 442-449.

(202) Kabsch, W. (2010) XDS, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 125-132.

(203) Collaborative Computational Project. (1994) The CCP4 Suite: Programs for Protein Crystallography, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 50*, 760-763.

(204) Steuber, H., Zentgraf, M., Gerlach, C., Sotriffer, C. A., Heine, A. and Klebe, G. (2006) Expect the Unexpected or Caveat for Drug Designers: Multiple Structure Determinations Using Aldose Reductase Crystals Treated under Varying Soaking and Co-Crystallisation Conditions, *J. Mol. Biol. 363*, 174-187.

(205) *MOE*, Version 2012, Chemical Computing Group, Montreal, Quebec, Canada.

(206) *Sybyl-X*, 2.0, Tripos International, St. Lois, MO, U.S.

(207) Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. and Richardson, D. C. (2010) MolProbity: All-Atom Structure Validation for Macromolecular Crystallography, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 66*, 12-21.

(208) Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. and Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Res. 28*, 235-242.

(209) Brautigam, C. *GUSSI*, 1.0.5k, The University of Texas Southwestern Medical Center, Dallas, TX.

(210) Schulze Wischeler, J. (2010) Carbonic Anhydrase II: A Model System for Artificial Copper Center Design, Protein-Guided Cyclodadditions, Tethering Screening and Fragment-Based Lead Discovery, Philipps-Universität Marburg.

(211) Lesburg, C. A., Christianson, D.W. (1995) X-Ray Crystallographic Studies of Engineered Hydrogen Bond Networks in a Protein-Zinc Binding Site, *J. Am. Chem. Soc. 117*, 6838-6844.

(212) Smart, O. S., Womack, T. O; Sharff, V, Flensburg, C., Keller, P., Paciorek, W., Vonrhein, C. and Bricogne., G. *Grade* (2014), Vol. 2014, Global Phasing Ltd., Cambridge.

(213) Laskowski, R. A., McArthur, M. W., Moss, D. S. and Thornton, J. M., (1993) PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures., *J. Appl. Crystallogr. 26*, 283-291.

(214) Chern, J.-W., Leu, Y.-L., Wang, S.-S., Jou, R., Lee, C.-F., Tsou, P.-C., Hsu, S.-C., Liaw, Y.-C. and Lin, H.-M. (1997) Synthesis and Cytotoxic Evaluation of Substituted Sulfonyl-N-Hydroxyguanidine Derivatives as Potential Antitumor Agents, *J. Med. Chem. 40*, 2276-2286.

(215) Moriyama, K., Nakamura, Y. and Togo, H. (2014) Oxidative Debenzylation of N-Benzyl Amides and O-Benzyl Ethers Using Alkali Metal Bromide, *Org. Lett. 16*, 3812-3815.

(216) Zheng, Y., Liu, B., Gou, Z., Li, Y., Zhang, X., Wang, Y., Yu, S., Li, Y. and Sun, D. (2015) Design of Novel CSA Analogues as Potential Safeners and Fungicides, *Bioorg. Med. Chem. Lett. 25*, 791-794.

(217) Sun, H.-X., Sun, Z.-H. and Wang, B. (2009) B-Alkyl Suzuki–Miyaura Cross-Coupling of Tri-N-Alkylboranes with Arylbromides Bearing Acidic Functions under Mild Non-Aqueous Conditions, *Tetrahedron Lett. 50*, 1596-1599.

(218) Imamura, Y., Higuchi, T., Otagiri, M., Nagumo, S. and Akita, H. (1994) Catalytic Properties of Carbonyl Reductase from Rabbit Kidney for Acetohexamide and Its Analogs, *Bioorg. Chem. 22*, 387-394.

(219) van Es, T., Backeberg, O. G. and Morrison, I. (1964) Some Sulphonamido-Derivatives of Stilbene and Acetophenone, *J. S. Afr. Chem. I. 17*, 95.

(220) de Benedetti, P. G., Iarossi, D., Menziani, M. C., Frassineti, C. and Benedetti, A. (1988) Multinuclear NMR and Vibrational Spectroscopy Studies of the Substituent Effects in Benzensulphonamide Inhibitors of the Enzyme Carbonic Anhydrase, *J. Mol. Struct. 175*, 37-42.

(221) Feichtinger, H. (1971) Notiz über die Willgerodt-Reaktion von 11-Chlor-Toluol-Sulfonsäure-(4)- und -(2)-Chlorid, *Chem. Ber. 104*, 1697-1698.

(222) Shil, A. K., Sharma, D., Guha, N. R. and Das, P. (2012) Solid Supported Pd(0): An Efficient Recyclable Heterogeneous Catalyst for Chemoselective Reduction of Nitroarenes, *Tetrahedron Lett. 53*, 4858-4861.

(223) Kleywegt, G. J., Zou J.-Y., Kjeldgaard, M. and Jones, T. A. (2001) MOLEMAN, In *Int. Tables Crystallogr.* (M. G. Rossmann, E. A., Ed.), 353– 356, Netherlands, Dordrecht.

# Eidesstattliche Erklärung

( gemäß §10 der Promotionsordnung )

Ich versichere, dass ich meine Dissertation

**Insights into Protein-Ligand Molecular Recognition: Thermodynamic, Kinetic and Structural Characterization of Inhibitor Binding to Aldose Reductase and Carbonic Anhydrase II**

selbständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe. Alle vollständig oder sinngemäß übernommenen Zitate sind als solche gekennzeichnet.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den _____ _____

(Chris Rechlin)